WEIGHING THE EVIDENCE:

VARIANT CLASSIFICATION & INTERPRETATION

IN PRECISION ONCOLOGY

U.S. Food and Drug Administration

Monday, January 29, 2018

8:30 a.m.

Food and Drug Administration

Center for Devices and Radiological Health

10903 New Hampshire Avenue

Building 31, Section A

Silver Spring, MD 20993

Reported by: Natalia Thomas

A P P E A R A N C E S

GIDEON BLUMENTHAL, M.D.
Deputy Director (acting), Office of Hematology and Oncology Products FDA/CDER/OHOP

REENA PHILIP, Ph.D
Division Director, Division of Molecular Genetics and Pathology FDA/CDRH/OIR/DMGP

JULIA BEAVER, MD.
FDA/CDER/OHOP/DOP1

MICHAEL BERGER, Ph.D
Memorial Sloan Kettering

JOHN DEEKEN, M.D.
Inova Translational Medicine Institute

VINCENT MILLER, M.D.
Foundation Medicine, Inc.

A P P E A R A N C E S (continued)

ANAND PATHAK, M.D., Ph.D
FDA/CDRH/OIR/DMGP

SASHIKANT KULKARNI, M.S., Ph.D, FACMG
Baylor College of Medicine

HOWARD MCLEOD, PHARMD, FCCP
Moffitt Cancer Center

APOSTOLIA-MARIA TSIMBERIDOU, M.D., Ph.D
MD Anderson

NEAL LINDEMAN, M.D.
Association for Molecular Pathology

HISANI MADISON, Ph.D, M.P.H.
FDA/CDRH/OIR/DMGP

HEIDI REHM, Ph.D, FACMG
Harvard University

A P P E A R A N C E S (continued)

KENNA SHAW, Ph.D
AACR Project GENIE

BEN PARK, M.D., Ph.D
John Hopkins Kimmel Cancer Center

KARLA R. BOWLES, Ph.D, FACMG
Myriad Genetic Laboratories

E. DAVID LITWACK, Ph.D
FDA/CDR/OIR

DANE DICKSON, M.D.
CureOne

DENNIS DEAN,, Ph.D
Seven Bridges Genomics

ROBERT GROSSMAN, Ph.D.
University of Chicago

A P P E A R A N C E S (continued)

J. LEONARD LICHTENFELD, M.D., MACP
American Cancer Society

ROBERT SCHUCK, Ph.D, Pharm D.
FDA/CDER/OCP

SUSAN MONAHAN
FDA White Oak Campus, Great Room, Building 31
10903 New Hampshire Avenue
Silver Spring, MD 20993
301-796-5671
susan.monohan@fda.hhs.gov
(On behalf of client)

C O N T E N T S

P R O C E E D I N G S

MS. MADISON: Good morning. How's everyone doing this morning, great. Thank you so much and welcome to "*Weighing the Evidence: Variant Classification and Interpretation in Precision Oncology.*" My name is Hisani Madison, I'm a Senior Reviewer in the Division of Molecular Genetics and Pathology in the Center for Devices and Radiological Health.

So before we get into the introduction this morning I'm going to give a few administrative FYI's. This will be recorded and is online in web-X so if you could please set your phones, computers and blackberries to silent mode, that will help reduce some of the background for those who are watching online.

Wi-Fi can be accessed in the Great Room using the code "public access" -- all lower case. There are food and beverages right outside for purchase in the kiosk right in the registration lobby and you could preorder your lunch boxes during the morning break if you like. You could also order them right before lunch.

Links for this archived webcast will be available on the workshop registration website shortly after the workshop but these slides will not be publicly available for you to download.

Also in about 45 days following the workshop there will be a transcript available on the same website as well.

After the introductions, we'll go over the meeting agenda but we have some printed copies right outside the door. Each session is set up to have multiple 15 minute presentations and a panel discussion which will be moderated by an FDA person here.

And after the moderator's section there will be about 10 minutes or so open for public discussion. And we encourage the audience to participate, to ask questions, as well as to continue the conversation online using via social media, using the hash tag "FDA Cancer Variants."

For the speakers, we will have a timekeeper right up front holding up slides for 5 minutes, 3 minutes and 1 minute remaining in your presentation. There's also a little red light, green light, yellow light here which will let you know when you have about 5 minutes remaining in your talk and as well when your time is over, so to stay on time if you could just keep an eye out for this little timer here.

And then next I would like to welcome Dr. Blumenthal, the Deputy Director of the Office of Hematology and Oncology Products from CDER to give us our first opening remarks, thank you.

DR. BLUMENTAL: Thanks Hisani and congratulations on putting this all together -- this great workshop. It's great to have the true precision oncology believers here on a Monday morning in late January.

So Hisani wanted me to say a few remarks and just give an update on the Oncology Center of Excellence so just to remind everybody the Oncology Center of Excellence -- this is around the one year anniversary.

It was founded in January, 2017 when FDA officially launched the OCE to leverage the combined skills of regulatory scientists and reviewers with expertise in drugs, biologics and devices to expedite the development of oncology medical products and support an integrated approach to the clinical evaluation of products for the treatment of cancer.

It's the first center of its kind at FDA to focus on a specific disease and I think some of the leadership at FDA has insinuated that perhaps if we get it right with the OCE, perhaps there will be other Centers of Excellence focusing on other therapeutic areas.

So just some of the highlights of the OCE in the past year -- we formed a Scientific Council to provide advice from the non-clinical perspective around the agencies on key initiates to pursue.

We formed disease-specific interest groups to discuss state of the science across various malignancies. On the approval side -- in 2017 we approved 16 new drug and biologic applications including the first -- we helped coordinate the clinical review for the first two cell-based gene therapies, the CAR T cells for refractory hematologic malignancies.

We also approved 30 supplemental drug and biologic applications. Several of these approvals also have companion or complimentary diagnostic approvals and Reena, I'm sure, undoubtedly will discuss some of this including several Oncopanels -- next gen sequencing platforms that were approved or cleared last year.

Also notable last year was the first so-called histology agnostic approval -- the PD-1 inhibitor pembrolizumab was approved for refractory MSI high solid tumors. This was sort of a landmark approval in that it was approved based on a biomarker rather than a specific site of origin and it opens up a plethora of interesting policy discussions including around biomarkers.

As we anticipate more histologic agnostic approvals, this underscores the need to get the biomarker testing right with appropriate standards for what constitutes biomarker positivity we're essentially in essence, redefining diseases.

With the increasing use of NGS platforms in oncology -- both for drug development and at the point-of-care to make treatment decisions, it's becoming increasingly important to discuss how to classify somatic genomic variations and how to interpret the results of these panels.

It's important to note that we won't get all the answers today, but we do want to understand the state of the science, learn from stakeholders on current best practices on varying classification, discuss use of public, private databases for classification interpretation and to discuss future directions including data sharing and harmonization.

While we made great progress, we know that we have a long way to go to reach the promise of precision oncology and it will need cooperation and input from all stakeholders. So with that I'll turn it over to Reena to give us a short update from the CDRH end.

DR. PHILIP: Good morning. Thank you all for coming to this workshop. We are looking forward to an exciting workshop. Our goal is get input from experts in oncology precision medicine on how to best weigh and evaluate evidence for classification and interpretation of sequencing results in precision oncology.

So in January, 2017 AMP ASCO and CAP published a joint consensus recommendation for standards and guidelines for the interpretation and reporting of sequence variants in cancer.

Still we understand the implementation of these recommendations is not consistently applied across all stakeholders. We also know that the multiplex tumorprofiling tests are reporting increasing number of variants day by day. so that gives uncertainty for the clinicians in the interpretation and prioritization of these variants with respect to their clinical significance and the optimal course of action.

So we are holding this public workshop to get input from experts like you to discuss how this genetic sequencing result is best implemented in patient management so that we can come up with innovative regulatory strategies to support the development of safe and effective precision-based drugs and devices for marketing.

As we all know the multiplex tumor profiling tests report many biomarkers. And these biomarkers may have a range of clinical evidence associated with them that are constantly changing as new science emerges.

At FDA we are committed to and work individually with the test developers to use the least burdensome approach for review of these tests. So last year a new approach was taken to the regulation of these tumor profiling NGS tests, incorporating the multiple levels of evidence -- or clinical evidence in our decision making.

A three-tiered approach for reporting biomarkers in tumor profiling NGS tests was taken. As you can see in this triangle the level 1 is the companion diagnostics tests - they are the tests that provide information that is essential for the safe and effective use of a corresponding therapeutic product, such as the drug.

So a tumor profiling NGS test may include companion diagnostic claims that are prescriptive for a specific therapeutic product as seen in the example here.

This is the lung cancer panel from Thermo Fisher Oncomine CDx Target Test that was approved in June last year. So this includes a companion diagnostic claim as you can see in the table here.

The tumor profiling NGS tests can also include biomarkers, cancer mutations with evidence of clinical significance, and the clinical validity of these biomarkers are established in professional guidelines but they may not be established with the test.

So those are the level 2 biomarkers and the level 3 biomarkers are the cancer mutations with potential clinical significance. The clinical validity of these cancer mutations are not demonstrated either in the professional guidelines or with the specific tests, but its suggestive based on clinical or biological evidence.

So this approach was taken when we de novo authorized MSK impact assay which includes only level 2 and level 3 biomarkers.

With this de novo authorization established a new class II regulatory pathway and so that means NGS tumor profiling tests are eligible for the 510K clearance process by applying to FDA directly or through an accredited third party reviewer like the New York State Department of Health.

And the intended use of the MSK impact assay as you can see here - it says the test is intended to provide information on somatic mutations -- in this case it was point mutations and small insertions and deletions and microsatellite instability for use by qualified healthcare professionals in accordance with professional guidelines and it’s not conclusive or prescriptive for labeled use of any specific therapeutic product.

We also approved FoundationOne CDx, F1CDx, November, 2017 and this is a broad panel and is a follow-on companion diagnostic for five tumor indications -- so it's a genomic profiling test of 324 genes, it also includes MSI and TMB in all solid tumors.

And this was also a breakthrough designated test and also went through the Parallel Review Program and I'm just giving you the links for all these -- what I described in my earlier slides.

So the three-tier approach is described in the first link and the SSEDs of Oncomine, F1CDx and decision summary of MSK are all on the public website.

With that I will turn it over to Hisani.

DR. MADISON: Thank you Gideon and Reena for giving us a great introduction and setting the stage for this morning's workshop. Again my name is Hisani Madison.

I also want to take the time to thank our public workshop planning committee -- they did a great job of bringing in some great speakers and thank you to our speakers for taking the time to come and speak with us today.

I want to give you guys a brief overview of the agenda which was touched on a bit by both Gideon and Reena this morning. We're going to start with Session 1 which is the overview of the state of science for sequence variant classification in oncology and its practical use in treating patients.

We'll have a quick break in between that session and then we'll go into the levels of evidence required for reporting variants and guiding patient treatment.

So in this session we'll talk a bit about the guideline paper that was published that Reena mentioned by ASCO, CAP and AMP.

And in Session 3 -- which is going to take place after lunch we'll be talking about the best practices for use of public and private databases, for variant classification and interpretation in oncology.

And finally we'll finish the day with Session 4 which is more forward thinking as we talk about future directions for data sharing, standardization as well as establishing some level of consistency in precision oncology.

I wanted to note for the speakers there is a little thing here that you can use to move your slides forward. So with that, I'm going to welcome Dr. Beaver, who will be the moderator for Session 1, thank you.

DR. BEAVER: Thanks Hisani and thanks everyone for coming. My name is Julia Beaver and we're really looking forward to today. The first session which I'll be moderating is, "*State of the science for sequence variant classification in oncology and its practical use in treating patients.*"

And I'll introduce our first speaker -- just logistically we'll have the three speakers present, then we'll convene for a panel discussion initially just moderated here and then I'll open it up to the floor for questions -- so we'll save the questions until the panel.

Our first speaker is Dr. Michael Berger, an Associate Director of the Marie Josee and Henry Kravis Center for Molecular Oncology at Memorial

Sloan Kettering Cancer Center.

He’s also an associate attending geneticist in the department of pathology and as a scientific director of clinical NGS in the molecular diagnostic service he oversees the development and bioinformatics associated with clinical sequencing assays.

He’ll be speaking on clinical sequencing and variant interpretation to guide patient treatment.

DR. BERGER: Great and thank you for the introduction and for the invitation, great. So I’ll be presenting the perspective of an academic cancer center -- Memorial Sloan Kettering, and how we perform molecular profiling and interpret the variants that we um, we see in patients with advanced solid tumors.

So the panel that we use as you already heard from Reena’s talk is the MSK impact panel -- like other panels of its kind, we collect tumor and in this case blood DNA and we prepare sequencing libraries, capture using probes designed to the most important regions of the genomic for understanding the clinical consequences of a patient's cancer and these results are analyzed and reported by a team within pathology consisting of bioinformaticians as well as molecular attending pathologists.

And what's notable about MSK impact are a few things. One is it's a matched tumor and normal assay and I'll describe the benefits of that later one. We sequence 468 genes which is a rather large panel for its kind but provides important information for other more complex signatures in the gene I'm going to also describe.

And we sequence a very deep coverage to ensure that we have very high sensitive for detecting low-frequency, either subclonal mutations or mutations in low purity tumors.

So the content of the panel is shown here. We have all the protein coating exons of 468 genes and this is meant to capture all of the genes with actionable mutations and cancer as well as targets of investigational agents in clinical trials that are ongoing or planned as well as additional genes that are frequently mutated in the cancer that may not be actionable clinical biomarkers today but allow us to collect population-scale data and integrate it with clinical outcome and drug response data to determine whether any of these mutations may have immediate clinical benefit as well as cancer susceptibility genes because we're sequencing matched normal DNA from patients.

We also have introns of recurring rearranged genes, some non-coding content like the TERT promoter as well as snips across the genome that allow us to perform better copy number assessment and other QC checks.

In addition to choosing the content of MSK impact we spent a long time optimizing the probe design for MSK impact to insure not just maximal depth of coverage but uniformity of coverage across targets to make sure that there aren't many exons that fall below our thresholds for commutations.

And this is compared to available whole exome kits and the graph that you can see on the right.

So we've been running MSK impact since 2014 when we received approval from New York State Department of Health to run as a clinical test.

All the testing is performed in the clinical environment and reported back to patients. As you heard we received FDA authorization for MSK impact late last year and in our molecular diagnostic service led by Marc Ladanyi we are sequencing and reporting out about 150 to 200 cases per week and you can see our progress since 2014 on the graph.

All together we've sequenced over 23,000 tumors from 21,000 patients to a mean sequence coverage of 720X. This is part of the impact team as you can imagine at an operation of this scale there are many, many people involved.

Ahmet Zehir leads the clinical bioinformatics team, Ryma Benayed leads the clinical next gen sequencing team within the department of pathology and molecular diagnostic service led by Marc Ladanyi.

So we've published an interim analysis of the results that we had compiled late last year and this is the distribution of tumor types that we had sequenced to that point.

And I show this to emphasize that this is being offered at Sloan Kettering across all types of solid tumors. This is representative of the distribution of cancer types that are treated in our center and it's not limited to just patients with lung cancer or colon cancer or melanoma where they may be FDA recognized biomarkers that are essentially required for genotyping.

We sequence patients with breast cancer and liver cancer and prostate cancer and brain cancer and many rare cancers as well. And this is important in identifying patients that may qualify for the basket clinical trials that have really emerged in the last couple of years where patients can be enrolled based on a molecular target independent of the histology of their tumor.

So I mentioned that it's a matched tumor and normal test and this has allowed us to not just query somatic mutations in a patient's tumor but also learn about inherited germline variants as well as identify mutations associated with clonal hematopoiesis.

So just to briefly describe our experience from what we've sequenced -- looking at somatic mutations we've been able to characterize the landscape of genomic alterations and more complex mutation signatures in patients with advanced cancer.

This was published last year and in our analysis we've seen that 13% of patients have been enrolled on genomically matched clinical trials on the basis of results that are obtained through MSK impact and I'll discuss that in a little bit more detail in a few slides.

And all of our results have been shared through the cBioPortal as part of our publications as well as the AACR GENIE Project that I also mentioned.

We have also since 2015 begun signing out and reporting germline variants associated with cancer predisposition so pathogenic and likely pathogenic variants are reviewed and signed out and recently published analysis for the first 1,000 patients to receive this analysis.

We have not performed this on over 5,000 patients and what we've found in a not so-unbiased cohort but more unbiased than the patients that are referred to clinical genetics based on family history and other criteria -- about 20% of the patients who received this analysis had pathogenic or likely pathogenic variants associated with cancer predisposition.

What was very interesting was about half of those patients had variants that wouldn't have been detected based on convention screening guidelines. They're either in tumor types or in demographics that wouldn't have otherwise received testing, or in genes that wouldn't have been considered for a particular tumor type.

Clonal hematopoiesis I won't say too much about but this is a phenomenon where mutations in hematopoietic cells lead to clonal expansion and can be precursors to hematological disorders or also cardiovascular disease -- this has been reported.

And through our analysis, because we're performing a deep sequencing on both tumor tissue and blood, we can identify low frequency mutations in blood that are absent from tumor tissue and attribute those to clonal hematopoiesis and we found associations with prior therapy, tobacco use, shorter survival and that this does confer increased risk of developing secondary hematological malignancies even though these are patients with solid tumors.

And while most of the clonal hematopoiesis is associated with general aging processes, a component is associated with prior therapy so these could lead to therapy induced leukemias.

So what I'm going to spend the rest of the time focusing on is what happens once we generate this data -- how do we interpret and disseminate and report these results -- so it's really this downstream component of our workflow.

All of our results are stored in a genomic variants database maintained by our department of pathology which we annotate using the OncoKB knowledge base which I'll describe in the next couple of slides.

In order to provide reports to doctors for their patients, facilitate clinical trial matching and allow for data mining and interpretation using the cBioPortal.

So there are many and you'll hear about many of these, I think, throughout the day. Many different knowledge basis for somatic mutations -- the clinical effects and clinical significance of somatic mutations when they're found in patients.

This is a slide prepared by Niki Schultz and Debyani Chakravarty who led our internal institutional effort to develop the OncoKB knowledge base which is shown at the bottom right.

And the way OncoKB works, like many others of its kind, are to annotate variants not just at the gene level, but for a specific variant within a specific tumor-type context and this uses databases, treatment guidelines, scientific literature, abstracts, FDA approvals, clinical trial resources to annotate the clinical significance of individual variants and individual genes for patients with different types of cancer.

And we use our own level of evidence system and I know there's a whole session devoted to levels of evidence so I'll save the details for that discussion, but suffice it to say variants are annotated according to whether they're FDA recognized biomarkers, whether they're standard care biomarkers associated with FDA approved therapies or investigational biomarkers for drugs in clinical trials and anything beyond that would be considered pre-clinical and research.

So variants get annotated at the alteration level as well as classes of variants like amplifications or all oncogenic or activating mutations and can be annotated according to the level that is appropriate.

And within our tiers of evidence, there's a distinction for whether the evidence is within the tumor type that the patient is presenting with or whether it's in other tumor types.

And these annotations go into the reports that we issue. This is an example of a lung cancer patient with a level 1 alteration which is an out fusion and a CDK4 amplification which is considered a 2-B by our criteria.

So in our published analysis from last year we annotated all the cases according to whether they had OncoKB level oncogenic mutations and all together, considering the FDA approves -- FDA recognized biomarkers to the investigation biomarkers, 37% of patients had at least one clinically relevant mutation.

And this is the number that's maybe low compared to other analyses that have been published but I want to emphasize that these are what our clinicians at Sloan Kettering consider to be actual determinants in the decisions that they're making so these are mutations that if found in a patient would have a significant impact on the treatment decisions for their patients.

We've integrated our molecular data now with the broader institutional database to facilitate the interpretation and matching to clinical trials so the molecular result -- the MSK impact to sequence results go into a database called Darwin which was developed by our medical informatics team that integrates that with surgical pathology results, other demographic and financial information, scheduling systems, other pharmacological databases and allows automated alerts to be sent to the oncologist who is leading the clinical trial as well as oncologists who are treating patients about the availability of a slot in a clinical trial for their patient.

So this is an example of one of those automated alerted actually sent -- generated by the system, sent from the PI who’s leading a clinical trial, one of the basket trials, to the oncologist who is treating the patient who now has a new mutation that qualifies them for that study.

And these can be timed actually with when the patient is due for their clinic visit. So actually, the most sophisticated searches return these results -- not when the mutation results are found, but a day or two before the patient is due for their next visit so it's not lost in the shuffle.

So this is a slide I wanted to spend a little bit of time on. This is our data from our first 10,000 patients as to how many of those patients were enrolled in a clinical trial based on specific targetable alteration found in their tumor by MSK impact.

And what we found was that 11% of patients at the time we performed this analysis had a -- had a trial match and this was based on almost 50 total genes -- some of which are shown here, including mutations in those genes, amplifications, deletions and fusions.

And actually this was an analysis performed in late 2016. When we repeated this analysis on the same cohort of patients 9 months later, the match rate went up to 13% which indicates that certain patients progressed on therapies, new trials opened, new knowledge emerged and so on.

What this doesn't include is um -- therapy, FDA approved therapies, that are administered on the basis of alterations that we find or off-label administration of therapies in other tumor types.

It does not include high mutation burdened patients or microsatellite instability patients who receive immunotherapy, patients where we identified germline alterations with clinical genetics follow-up.

So there are a lot of ways -- and also information that helped clarify or change diagnoses. So there are a lot of ways that this information is used by clinicians at our center.

All of the information is -- all the results are shared internally at our center on a daily basis in the cBioPortal and we've been releasing these publicly in batches to the community.

I'm sorry that URL didn't show up, but if you go to cBioPortal.com/ -- I'm sorry, cBioPortal.org/msk/impact you can access the mutations that are clinicians see but they would be in an obviously de-identified HIPAA compliant manner -- and also we've been sharing these with the AACR GENIE Project where now I believe over 30,000 patient profiles across the 8 initial institutions have been shared.

And this allows investigation into rare alleles, rare tumor types and the kind of things that a single institution wouldn't necessarily be able to build a sufficient caseload to study.

And just one last point -- I've alluded to this throughout the talk, but in addition to individual alterations that receive their own annotation and curation, larger panels like MSK impact can reveal complex, clinically relevant genomic features including tumor mutation burden.

As we know different tumor types tend to have different average numbers of mutations and within a tumor type there's often a broad range sometimes associated with environmental exposures like cigarette smoke or UV exposure or intrinsic genetic defects like microsatellite instability.

Specifically, with regard to microsatellite instability we've been running a bioinformatics tool developed at Washington University called MSI

Sensor which allows us to identify a microsatellite instability signature -- not just in patients with colon cancer and endometrial cancer where IHC testing and MMR, PCR, MSI, PCR are common -- but we’ve observed MSI signatures in a large number of tumor types which has allowed patients to go on to receive immunotherapy through clinical trials.

So to summarize at our institution we’re using targeted and NGS panels to reveal many different types of clinically relevant mutations, point mutations, copy number gains and losses, rearrangements and mutational signatures.

For us, large scale implementation of clinical sequencing has been feasible and almost necessary in order to um, optimize the treatment decisions that are being made with regard to available therapies, clinical trials, diagnostic decisions and otherwise.

We’ve been sharing data because this is the best way to enable large scale biomarker discovery and to characterize and study rare tumor types and as you'll hear throughout the day, variant annotation is necessary to inform treatment selection and matching patients to clinical trials.

So this is part of the team. Our team at Sloan Kettering who have been invested in this effort for many, many years now -- I'd like to highlight Marc Ladanyi, Maria Arcilla, Ryma Benayed and Ahmet Zehir in Molecular Diagnostics; Jose Baselga, David Hyman, David Solit for Institutional Leadership; David Klimstra for Department Chair of Pathology and Niki Shultz and his team who've done most of the work in developing the OncoKB database in partnership with clinical fellows, attendings and research fellows throughout our institution.

So thanks very much and now I guess I'll take question from the panel.

DR. BEAVER: Thanks so much. So our next speaker is Dr. John Deeken, who is Chief Operating Officer of the Inova Translational Medicine Institute. He's also a practicing medical oncologist and Senior Vice-President for the Inova Health System with a clinical focus on the treatment of patients with head and neck cancer.

Additionally, he's Associate Professor at Virginia Commonwealth University and he will be speaking about tumor profiling at a community hospital system.

DR. DEEKEN: So that is a tough act to follow. So let me tell you about maybe the other end of American healthcare and where a community hospital system in northern Virginia has been trying to keep up with the great science and the great access for patients in terms of tumor understanding, tumor profiling and targeted therapies.

Inova -- and some of you know who live in the area, is a hospital system in northern Virginia, a 5-hospital system and with almost 2,000 beds and about 4,000 -- 400,000 ER visits per year.

We serve the northern Virginia area. Our direct catchment area is about 2.3 million people, the larger catchment area is about 6.5 million people if you include the region. And that's the care area that we provide -- so not a small health system, it's a non-profit health system.

And the leadership a few years ago decided that a number of its strategic goals was going to be personalized medicine -- precision medicine and trying to incorporate -- as many systems are in the country, the understanding of how genomic medicine can improve care and improve prediction of illness as well as better care.

The system, as part of that, decided to invest in a large effort to build our own internal capabilities, our lab capabilities and recruited Dr. John Niederhuber when he left directorship of the NCI and came to Inova in 2010 to create our genomic focused research institute.

And I tell this because this is what we developed into our cancer tumor testing platform. The overall goal of ITMI was to pursue genomic research and how it can inform best practices in medicine.

In 2015 we became CLIA certified. Last year we became CAP certified. Our staffing includes lab personnel as well as clinical and bioinformatics research staff -- so that's the setting for what we could do in terms of building up our capabilities.

Um, at IT we've been also at the same time in our cancer center recruited some top leadership from around the country -- Dr. Skip Trump is our cancer center director, he came from Roswell Park.

Dr. Joan Schiller came from UT Southwestern to develop a core group of medic oncologists that have a quasi-academic focus, they're not just busy clinicians in the community, but also ones that are developing research, clinical trials and new efforts, including our molecular tumor board.

So we created about a year and a half ago for refractory cancer patients, the ability to have their tumors tested to look for other options if standard of care options had failed for them.

Numerous systems and academic cancer centers have created this type of molecular tumor board process but that was our attempt to offer for patients who didn’t have treatment options, a way of having additional treatments identified.

Along the way, given the rising need to have targeted therapy testing, we created again in house assays typically for a variety of tests including ETFR, RAS, BRAF, microsatellite instability, brain tumor methylation and MGMT.

But again our main focus was to look at traumatic tumor profiling using NGS to support our molecular tumor board. This was the platform we used, and I apologize -- it says Illumina, obviously the Oncomine panel is Thermo Fisher that we run on the ion torrent machine.

We initially used the hotspot panel which had about 50 genes and targeted mutations in those 50 genes which are known to be oncogenic drivers. Then we moved towards the Oncomine comprehensive panel which is much more comprehensive in terms of hotspot genes, full gene coverage as well as copy number variants and fusions to look for -- to offer as our testing platform.

Since we didn’t have the infrastructure of a great cancer center and a basic science faculty to help us with interpretation, we had to bring that in from the outside.

So the way we developed our workflow for this testing was we did the testing in house with our own ion torrent as well -- and again off the shelve, Oncomine testing platforms.

We developed the raw data, we developed a collaboration with a Washington University company called Pierian DX which has the knowledge base and curates the literature for updates in terms of identifying variants of significance that can be used as clinically actionable.

They were turned back to us based on the interpretation of the raw data -- a tumor report, that identifies actionable mutations as well as variants of unknown significance.

They also can identify clinical trials that that patient might be eligible for and then from our molecular tumor board again, since we don’t have a basic science faculty around or esteemed pathologists that can help identify --

help us identify the variants that might be the best match for patients, we actually again came up with a commercial relationship with N of One and their experts actually call in and participate in our molecular tumor board every week to help us sort through the data and the findings on each individual patient.

As we all know when you run a tumor test on a patient you can come up with numerous mutations so you are trying to find out what are the most relevant, the most actionable, the most likely to be driver mutations in that patient as opposed to passenger and to best determine the right therapy for the patient.

So again since we didn't have that expertise as in depth as we'd hoped, we partnered with N of One to create that.

We've run about -- not 23,000 but about 200 patients through this process over the last year and a half and these are recent reports that we presented last year and also at GIS just last week. And again we'll talk about evidence at the next session.

But our patients we had a pretty significant numbers that had tier 1 level evidence as well as tier 2. Not surprising, we had a number of hot spot mutations that were found in patients -- especially P53.

We use this process to identify patients like at MSK to be eligible for clinical trials -- so we have a number of basket trials open at our institution including NCI match, including ASCO's TAPUR and some other single drug basket trials.

We tried to use this process to identify the right patients for those studies. We too, found about a 10% match rate -- only about 10% of patients had tumors, mutations that led them to be eligible for those trials unfortunately.

We also used this to find compassionate use options for patients that might be eligible for a targeted therapy that was already approved but not indicated for their tumor type and they've had decent success getting access to those drugs for patients who had mutations that again -- there was a targeted therapy in the market for but that was not in one of these clinical trials.

Um, in our real world the big challenge has been finances. A cost of that platform typically these platforms, is a little under $2,000 -- that's not counting the overhead, the lab tech, the N of One, the Pierian DX cost.

When we spent a lot of time in 2016 looking at payer coverage for this and wanting to cover NGS coverage, we only had 1 of 10 payers who were willing to pay anything.

They actually paid sufficiently that if we had -- if all of our patients had their insurance we'd be doing alright but since a very few of our patients actually had that payer coverage, the ability to pay for this from insurance was, was minimal.

So instead we pursued a philanthropy approach -- so we had a large philanthropy effort to try to raise the funds to support a molecular tumor board, including this tumor testing, but after that sort of effort went as far as it could actually -- for 2018 we decided to take this out of in-house and move it to foundation medicine given the recent approval and their better success in terms of financing this sort of testing.

So that's -- we tried for a year and a half to support this kind of testing at our institution and again, with philanthropy support was able to do it but unfortunately the real world of payers has not quite caught up with where the science is right now and unfortunately we actually moved that large profiling effort to out of house.

Our plans for this year is to look at more targeted panels, payers and to an extent CMS does better in terms of covering that so the new panels that are covering for lung cancer, colorectal cancer and others is now our new focus in terms of NGS profiling.

We're also continuing to develop single gene tests, either as the companion diagnostic or an in-house lab-directed assay that needs to be done as drugs get approved for those for like for example, midostaurin for FLT3 -- when that got approved we knew for our hematological oncologist we needed that assay to be able to be done in 24 hours for a newly diagnosed AML patients.

And interesting -- and again in the real world or the community world something to be aware of -- one of the reasons why this made sense to a health system like ours and many, to have it in house was because if tumors were tested in the in-patient setting or within 14 days under CMS guidelines, on an in-patient admission the testing wasn't paid for -- it was all covered by the DRG payment that you got as a lump sum payment for that patient's in-patient stay.

That meant that often times pathologists would hold on to those tumors and send them out on day 15 or whenever they got around to it to get the testing done. And if you think about the turn-around time oftentimes of that being sent out is 2 to 4 weeks or even longer, the treatment decisions on that patient can be delayed significantly by that added time flow in terms of that.

So because of that and because oftentimes clinicians were pushing our hospital to do it quickly, the hospital was eating that cost and it made sense for us to move that in-house because we could do it cheaper and the hospital was out that money anyway.

This past fall -- last summer CMS proposed a change to the 14 day rule and actually in the fall they did change it. So we're still looking at the economics of does that change our financial calculation that doing these tests in house saves money to the hospital since they -- if it got sent out they'd have to pay to those outside testing companies to do.

So we're still evaluating where that changes that pivot point to what we do in house versus what we do out of house and we don't have good answers on that yet for what we're doing.

So that's actually the end of my talk and thank you very much.

DR. BEAVER: Thank you. So our last speaker is Dr. Miller who is the Chief Medical Officer of Foundation Medicine and previously an attending physician at Memorial Sloan Kettering Cancer Center.

He will be speaking on his industry perspective of variant classification.

DR. MILLER: Thanks Julia and thanks Hisani for inviting me. Amazingly -- I don't think there's any redundancy with the first two talks. That's a -- I'm sure that was in the planning session right and I think our prior speakers have set the table well for the points I wanted to hit.

Really as you -- you know I'm humbled to be part of the team that um, was successful in the parallel review process working with FDA and leadership and CMS for the approval that Dr. Philip alluded to earlier.

And really there were two key drivers of that. One is -- and I think everyone has alluded to the fact the need for excellence as therapies become more binary in their ability to parse patients.

You have the marker -- there's a really high change you'll respond. If you don't have the marker, with many drugs there's a very low chance of response. And this is in contra distinction for those who aren't in the oncology space to decades of treating patients with cytotoxic chemotherapy where people would talk about putative IHC markers -- well if you had one positive you might have a 32% chance of response and if you were negative a 21% chance of response.

Those don't make big differences in clinical care of our patients in general and really what they do is perhaps drive where one might choose chemo A or B or what's used first line and second line as opposed to third line and fourth line.

So in late 2017 -- and the second part, what was the second driver is getting paid. So um, we actually just had deeper pockets maybe than the Inova system and a lot of investors who have faith in our belief.

It isn't that we have more money per se and as was -- as our CEO would say we are a pre-profit company meaning that we're still doing many, many tests because of the right thing to do without ultimately getting paid.

And so the second driver here was that as arduous a path perhaps, as parallel review was -- although it was always fair and responsive and a great relationship -- at the end of the day there was someone that if you got, you know, through this process there was a path to get paid if the can could not be kicked down the road forever.

And so um, there's one thing many in the room should be united on even if it's tangential perhaps, is getting paid for these tests. So, FoundationOne CDX and this is not the exact label so no one gets chest paid, is a -- this is our, sort of abbreviated statement is the next generation sequencing base in vitro diagnostic for detecting the four classes of DNA-based alterations, the short variants, copy number changes in 324 genes as well as deletions and select gene rearrangements as well as two genomic signatures -- MSI and tumor mutational burden using DNA isolated from clinically available specimens, FFPE based specimens.

And the intended use statement can be found at that URL there. So things we'll hit on today and some of the talks do bleed into other sessions but I think in some ways that may be helpful because it gives different perspectives.

My perspective will be that of a medical oncologist largely and so I think that's important in the sense that the space -- and most of us are talking about metastatic cancer as much of the content today -- that's really different than a lot of other disease states.

And most of those diseases in the advanced state are incurable. Treatments are variable efficacy, unambiguously getting better but they're still diseases where a doctor wants just a sniff of a therapeutic option for his or her patient -- GBM, pancreatic cancer, et cetera.

Breast cancer -- maybe it's a little bit different. So that's clearly a nuance that I think is unique to this discussion and at some level does touch the disease ontologies but also as you know we sort of have taken this tissue agnostic approach at um -- in the way we frame many of our discussions and reports.

So -- role of subject matter expertise in variant interpretation; interpretation for clinical decision making and in that context and then the biggest challenge of course reporting and nuancing -- providing language around non-canonical but clearly clinically relevant findings.

So there are three sections to the new report some of you will see. The FoundationOne CDx report -- the first is the FDA approved content. The second is a professional services section and the third are the appendices.

All genomic findings outside of our FDA approved claims will be shown in the box on -- below what I -- below the report, to the left side of the slide and on the report below the orange section and the genomic signatures for MSI and TMB are included with every test with results on page 1 and then the interpretative context is providing the professional services.

The FDA approved CDx or companion diagnostic claims are shown with the associated therapies, listed in alphabetical order by brand name with generic name included for quick recognition.

We have our agnostic as far as biopharma partner despite various investors, we have partnership with three dozen folks, we're looking for more -- it's a -- it's a, I think -- and that's one thing that we're completely -- it's part of the reason I took the job at FMI is because I didn't want to be linked to a specific drug, a specific age and a specific path. I wanted it to level the playing field.

So what about the interpretative context and this is really the crux of the challenge to some extent -- a lot of what the discussion is about today that's providing professional services.

We classically have divided um, our reporting in this section into three groups -- therapies with clinical benefit in the patient's tumor type, therapies with clinical benefit in another tumor type and then clinical trials either directly seeking net variant or mechanistically tied to that variant or an alternation similar and again -- another crux of the discussion for that particular tumor.

So what's involved in this interpretative process? And again this is really sort of the -- getting to the more challenging issues under discussion today and certainly one of the things that I think FDA and CDRH have been incredibly thoughtful about is the pace of change in the field and the need to be able to move logically and quickly and not lock things down.

Because by the time we all leave today, we may need to iterate something we all do on reports. So of course analyzing genomic alterations, what's the impact on DNA protein and function, the impact and molecular and cellular pathways and of course the evidence around these can be incredibly varied.

At Foundation Medicine you know, we have a large number -- I don't know the exact number, it's a lot of people of very-well trained bioinformaticians who confirm the variants, make sure they're not artifactual, et cetera -- it's sort of the human overlay on what comes off the sequencer constantly for our non-regulated products, you know, working on improving algorithms and so forth.

The second component though is compiling and interpreting the evidence. And this is, you know, drawn from multiple sources -- scientific publication at conferences, medical experts, online databases.

In this area we have a team of doctoral level trained scientists who are constantly scouring scientific publications, proceedings, abstracts, posters and trying to put the most balanced and current context to what we report.

And I preach to the team you know, taking the hat of the medical oncologist generally when we do this, to err on the side of providing more options but at the same time being very precise in what information we -- and what references, how did we get to providing that information so that even if we differ from someone else, we can tell you how we got there and you can tell us how you got there and then we can compare and contrast and the doctor can make the decision appropriately.

And then finally, of course, is summarizing therapeutic implications and gene interpretation and getting a report back that the medical oncologist probably has.

I think the average visit in U.S. oncology for a patient is about 7 minutes so the doctor has very little time to take that information and integrate it into clinical care.

So many sources -- these have been alluded to in part, that are used in this variant interpretation process. I would add to this, you know, that we have the benefit of having profiled many, many patients since our first test launched at ASCO 2012 and therefore our database foundation core, which is largely run on the same platform from 2000 -- you know, from that time.

And so there's a great degree of rigor and homogeneity there -- can also play into this. And we've been big proponents of sharing genomic data

-- having made data on a couple of thousand pediatric patients available and then more recently the largest single contribution of which I'm aware to GDC of 18,000 cases on a prior bait set which bookends nicely with the data from TCGA in that dataset.

Many folks required -- whoops, in this genomic interpretation process and unlike maybe Memorial where Dr. Berger can get on the phone with Dr. Baselga and say, "What's the deal with this breast cancer, could it really be your mark, but now you couldn't say could it really be this, this or this because genomic lead has these features.

In general -- this is a unidirectional flow of information from us back to the doctor absent the richness of clinical context that we would desire -- although we certainly, with proper guardrails in place do seek to gain as much information as possible in cases where we see things that don't make intuitive sense to us.

But many teams and folks and steps along the way -- and for those of you who have tried to build an assay and people often ask what's proprietary, this and that -- and while there may be proprietary components, the key pieces you have 8 or 10 different steps, all of which need to function at 99% plus for your assay to be successful because on the end of that -- the doctor will only see whether or not he got a report back in a clinically relevant timeframe for the patient, who's returning for a clinic visit.

So, to summarize and to get as many different types of information that a doctor would find important for treatment decisions and clinical and pre-clinical data unfortunately is often unclear, complex or conflicting.

And we should acknowledge that right? We've all turned away -- well some turned away papers because maybe they didn't reach the same conclusion as their own work but often times there are things in the literature that are contradictory and difficult to reconcile, even among excellent groups.

The sources for varied pathogenicity may conflict and I think some of you in the audience reached out to us to help improve some of our things but also vice-versa.

Treatment guidelines don't cover all cases -- for example rare tumor types, recent findings, contraindications, uncommon variants and diseases are the -- often the norm rather than the exception when one does this type of broad-based, unbiased assay that studies hundreds of genes and sequence the entire coding sequence of those genes.

And depending on context I alluded to earlier only a type of information needed to support clinical decision-making may vary. It takes subject matter experts with strong scientific and biomedical backgrounds to develop the most balanced interpretative product.

And this is just an example of a few things. Any medical oncologist would want to know if his or her patient's tumor harbored one of these alterations -- it's 99% plus. And yet, depending on where one draws the levels of evidence and they may or may not be in a salient place on a particular report.

I'd also point out when we get to the guidelines piece it's a little bit of a slippery slope. Guideline committees' can have the world's best surgeons on them who have no training in molecular pathology and who are used to things that you know, we either are 100% or zero -- meaning that if a drug helps 17% of the patients that may or may not be relevant.

So guideline committees are only as good as the expertise of the committee members. And if you look at a body like NCCN which does great work, they are solely diseased ontology focused now -- there's not a biomarker compendium like there is an antiemetic compendium or there is a growth factor support compendium but maybe there should be, you know, going across tumor types.

What about professional guidelines beyond that? Which one should be followed? Should the manufacturer decide what should be considered? What if they're contradictory? And of course, many cases in which the guidelines are not straight-forward.

And parenthetically for an institute maybe like Inova, there is a feat of these things -- they're not cheap in some places, particularly if you need to have multiple subscriptions.

So I alluded to this earlier and there's actually a duplicate row here. Currently we have, you know, our caveating of reporting is you know, and this is sort of A trumps B, B trumps C, C trumps D.

Approved therapy is indicated by FDC in the tumor type and then approved therapies as indicated by guidelines in the tumor type, approved therapies outside of the indication but supported by clinical data in the patient's tumor type -- approved therapies outside of indication supported by pre-clinical data and clinical trial relevance -- and, this is where I think it is a bit of a challenge too and something perhaps to discuss.

I don't feel any of us are -- I don't feel I'm in position to judge whether or not a new therapy targeting RB loss really should be listed, you know, tied to a particular therapeutic trial or not -- maybe this will be the one that works, or maybe it won't be.

We're doing a heck of a lot better in biopharma. So to make this distinction based you, you know, cell line data, grab data, whatever it may be that's all there is -- we're not going to list this, have it buried on a report and not have trials prominent for that patient is I think also a bit of a slippery slope into privacy of options and options are what our patients need.

So I'll stop there and thank you and I think it was only like 48 seconds over or so, that's good, thank you.

DR. BEAVER: So we'll invite our speakers and panelists up to the stage and I'll introduce our panelists. We will be joined by Dr. Donna Roscoe, who's the Branch Chief of the molecular genetics branch of molecular genetics and pathology at FDA CDRH. So perhaps I'll start with a question for

Dr. Roscoe.

We heard Reena talk earlier and heard some of this um, alluded to, but do you have anything you'd like to add to describe FDA's perspective on diagnostics and/or variant classification in this space?

DR. ROSCOE: I think --

DR. BEAVER: Oh, you got unplugged.

DR. ROSCOE: I think Reena did a great job describing our three-tiered approach. Ultimately what we were designed to do, what we strived to do is have validated tests on the market which allow for dynamic use within the clinical setting which enable clinicians to optimize patient decisions and that was what the MSK authorization was designed to do -- to allow biomarkers to be fluid within their various claims while simultaneously approving the foundation medicine for very specific evidence that they support companion diagnostic use.

And so we're always -- the purpose of this meeting is that we're always trying to gather the state of the art in terms of what is the practice, what is the most beneficial route to getting accessibly, analytically and clinically validated test to market, but ultimately we are an organization that's interested in evidence so how can we assure that physicians and patients are getting state of the art evidence and the most appropriate evidence at the time?

So hopefully we'll get that information from this workshop.

DR. BEAVER: Thanks and then this was a question I sort of prepared the panel for on our planning calls but one of the goals of today is to really get feedback from stakeholders about how FDA can improve and be involved in this discussion.

And so I'd like to actually go down the panel. We can start with Dr. Miller but can you give us um, your thinking on what the role you'd like FDA to play in variant classification and how do you see FDA as being helpful or unhelpful in moving this field forward?

DR. MILLER: Well I think to date um -- probably one of the most helpful ways is to almost by definition if we look at where we are now and put something into place we're going to be behind the times.

So we need to wear our respective caps of what will drug approvals look like -- not a specific agent but classes, therapies, indications, labels, 6 months or a year or 18 months from now.

So I've certainly been impressed by robust data with track inhibitors in tumors containing track 3 or other fusions and those alterations are so uncommon or rare that I would think there would be an opportunity for another tissue agnostic approval, now that's just my medical oncologist hat.

So with that being said that won't be -- that's not the first and it's probably not going to be the last so how do we think about that as far as biomarker testing? Because if one did have a track inhibitor approved in let's say pan cancer the doctor may still say, "Well how many patients were treated and included with my -- I'll just make it up, adenocarcinoma of the prostate and that sort of thing."

So um, I think our challenge is to almost work from the framework of where we believe the field will be globally in 12 or 18 months in thinking about both the evidence piece and variant classification because there needs to be a new way to think about evidence by definition in precision medicine -- we're not going to have randomized trials commonly, we're not going to have several hundred patients and tumor types where they may only be several hundred patients a year.

So how do we think about that? And certainly, I know there's some discussion later about ways to collect data in some of those tumor types and what that might look like -- will that get us where we need to go?

DR. BEAVER: Thanks.

DR. DEEKEN: I would agree with everything that Dr. Miller said. The one thing I would sort of add I know I focused on cost and sort of painted a sour picture at the end about our effort based on cost but I think as we are looking at the cost of whole genomic sequencing and the rapid decline and the cost to do that, I think we're also seeing and should see -- I hope, dramatic reductions in costs in terms of tumor profiling.

So therefore, if it's going to be as cheap or expensive to do NGS panel on a patient's tumor as it is to do two or three RAS and RAF mutations, the focus on single gene tests, and companion diagnostics for a specific drug might be less critical as knowing the underlying infrastructure that we treating physicians are going to be using in terms of tumor profile testing and then matching the right drug to that patient that's approved or on a clinical study.

So the focus on companion diagnostics one-offs might be less relevant moving forward if the cost curve continues to bend the way it should and the likelihood in 5 years or not too farther after that every patient will have the benefit of what Sloan Kettering has in terms of profiling when they walk in the door and looking for options then and down the road.

So that would be in terms of focus on priorities -- I'd be ready for that future because I think it's here in some places and coming elsewhere soon enough.

DR. BERGER: Yeah, I'm not sure I have much else to add. I agree with everything that's been said. I mean I guess I would emphasize that with respect to variant curation and classification it's -- it moves, sorry, it moves very quickly.

I think I would echo what Dr. Miller said that we lock down what we know today it's going to look very different 6 to 12 months from now so we have to have frameworks for interpretation that recognize the dynamic nature of the information that we have, sorry -- am I the only one hearing this? (microphone feedback)

I'm getting -- okay sorry. And yeah, I mean I think the expertise in this area is, is spread very broadly. I don't think any single person or center or committee or guideline's group can really speak accurately and comprehensively about the clinical significance and mutations.

Papers that have been published may be discredited. New studies may be well underway and that information may not be as broadly available but we wanted to make sure that when reports are issued that patients benefit from everything that the community know and has proven with evidence.

DR. BEAVER: Thank you. So we've touched on um, costs in the parallel review process a bit but um, Dr. Roscoe if you could just provide your FDA perspective on how that process works and any comments related to that.

I think we've touched on it but perhaps not yet described what parallel review is or entails.

DR. ROSCOE: Okay I think actually we were hoping to not touch into the basis of parallel review, but so I'll let foundation 1 discuss that about their experience with that, but ultimately we've only done it twice.

Notably we've only done it once with the Exact Cologuard test and now with the Foundation

Medicine test, it's not meant to be taken lightly. CMS takes this very seriously so I'll let Dr. Miller talk about that.

DR. MILLER: I will stick with the topics at hand and not spend a lot of time except to say that I provided the rationale for why we chose that path and in part sadly some of it was driven by the need to hopefully reliably get paid for some of what we do.

And we found it hard but balanced there in response of pathway from all involved.

DR. BEAVER: Thanks. So I've seen this in my second opinion clinic that I have but we do see panels come back to us on the same patient on the same specimen with differing reports, different variant calls, different recommendations and so my first question would be to Dr. Berger why does that happen, in general?

DR. BERGER: Right. I think they're many reasons. Obviously different panels have different content so some genes may be sequenced in one but not the other.

But I think more significantly than that is, you know, there are certain types of test sequence -- only tumor DNA and others match to normal DNA so what might be reported as a variant in a tumor only test may have been appropriately filtered out as a germline variant when matched normal DNA was sequenced.

So I think that's a big different that can lead to discordant results. Different tests have different thresholds for detection sensitivity. One tests might call down to mutations only in 10% of DNA molecules where another might be powered to detect down to 5% or 2%.

And then of course, different tests may have different criteria for what makes it into the report. Some tests may report all variants that are detected or all somatic mutations that are detected within the panel.

Others may limit the reporting to those that are deemed to have clinical significance. So I think there's a whole number of reasons why mutations may vary.

I think the concordance -- you know, despite reports that are unpublished, the concordance is generally much higher within the clinically validated or accepted biomarkers but um, but certainly there may be discordance there.

And it also depends on what actual tissue was sequenced. There may be separate sites of metastases or a primary tumor versus a metastatic tumor that may have genetic heterogeneity.

So different tests may have been run on different samples, so there are many valid technical reasons why that might occur and typically the actionable mutations are more concordant than others but that's certainly not a blanket statement.

DR. BEAVER: Did you have something to add?

DR. MILLER: I would just add -- and this is in part, even assuming all tests are performing optimally or 100% for what they do, that the details around what a given assay finds or does not find is essential then because when we did our for example -- a couple of our papers on tests that were negative for a certain EGFR alterations on prior reporting and you go back and look at the source document -- one still can't find out what EGFR mutations were exactly tested for.

So is this an error of omission or commission so to speak, so that's another piece to the puzzle. Then of course certain platforms -- even if you're doing a superb job with them are less able to detect certain classes of alterations.

Effusion detection of course is more challenging you know, as is insertion deletions and base subs where some of the publications around concordance have been -- that's sort of the low hanging fruit.

DR. BEAVER: Okay and Dr. Deeken, how clinically do you handle that sort of discordance or how might you handle that?

DR. DEEKEN: I think that's a tough question and I would say we often don't have the benefit of two tests to sort through so I think we cross our fingers and hope that everyone got it right.

I think to the point of Dr. Berger I think a key question that we’re facing in clinical trials as well as in standard of care is what do you biopsy?

Can you use archive paraffin, do you need a new biopsy to do that on and run it at that time in terms of the best next treatment for patients I think that -- I think the evidence is saying we need new biopsies but that puts patients at risk in terms of the risks of performing new biopsies on patients.

And our great hope of plasm markers is not panning out in terms of genomic science so I think um, I think we’re left on the clinical side hoping the experts got it right as best we can knowing it’s not perfect and hoping that it gets better in the years ahead.

But I think that’s a standard problem that I think we just swallow hard and try to keep going with.

DR. BEAVER: Thanks, and we heard a little bit about foundations ability to curate or update different variants. Dr. Berger, how does the MSK impact at your institution -- how do you update information regarding your panel for instance, of U.S. that now has enough evidence to be called deleterious and how do you determine how much -- I know we're going to touch on some of these topics later today, but just general thoughts on that.

DR. BERGER: Right, so thanks for the question. So our annotation all comes from OncoKB knowledge base that I described and that is led additional by our -- or is led primarily by an informatics team who developed the structure for it as well as some full-time staff curators, but most of the curations and updates come from, you know, a set of fellows and clinicians across the different disease teams who provide their disease-specific expertise.

So it is constantly being updated. Now for us there are two ways that physicians interact with the molecular reports and annotations. One is with the report that's issued by molecular pathology when the test is performed and that remains static.

If there's some critical update or a new type of alteration with clinical implications that our bioinformatics pipeline becomes able to detect, there may be an addendum or an amendment that's issued to the report.

But we don't update those reports when new knowledge emerges that might reclassify something from a VUS to a clinically significant mutation.

But the other way that physicians interact with the results at our institution is through the cBioPortal. I showed a screenshot of that -- it's a website that's updated daily with the new MSK impact results that are -- that are delivered and the annotations from OncoKB are always displayed in real time.

So there's actually a link from within the medical record directly to the patients results in the cBioPortal that we added about a year ago when we found out that oncologists were actually searching for their patient's data in the portal blindly when there was no link, or there were no identifiers, but just based on the specific spectrum of mutations that were found -- they found this a really useful tool for understanding the most up to date information about these variants that are being detected, both from a clinical standpoint and which mutations are recurrent, which ones are known hotspots, which are emerging from meta-analyses across TCGA and our datasets and other big datasets.

So there's been some divergence in some cases. If you go back to the original report it has the original annotations but if you were to follow that link to the cBioPortal and the website that contains them, then you're getting the most up to date information.

So that's what I think oncologists -- or we're encouraging oncologists to do as they're continually trying to find what's best for their patients.

DR. BEAVER: Thanks and yes?

DR. MILLER: If I could just weigh in. An important topic that Dr. Berger referred to that is crucial to this I think is the difference between germline and somatic and oftentimes the tumor profiling can find germline variants and one needs to have an operation in place or at least an understanding with the clinicians that if it might be germline, you need to further that patient's evaluation for the betterment of their family.

And oftentimes that's forgotten in this piece. And the way to detect that is by having mass germline and somatic but oftentimes I think we practicing clinicians are thinking about the drug for the patient in front of us but oftentimes it can be -- if it's a germline that's crucial information that needs to be followed up with genetic counseling in that kind of operation and that, I think, is falling behind in terms as the science pushes clinical care forward for the cancer patient.

DR. BEAVER: Thanks and in thinking about the report from these tests, how critical would you say it is to report the allelic fraction of the mutations?

For instance, thinking that a patient who has a less than 5% allelic fraction of a certain mutation may respond differently to a targeted agent or a clinical trial than someone with a 20% allelic fraction.

And we could start with -- I'd be interested actually in all of your thoughts so maybe start with Dr. Miller and come back towards me down the line.

DR. MILLER: Well I think the -- as some of you know we do TCGA work, we generally work with specimens that were 60 or 70, 80 or more percent tumor content and there we had the wherewithal to saw alright we won't study this one it's 32% tumor and our algorithms were making these calls have not been matured to that -- to that level.

But of course, in clinical practice many of the specimens we receive when we do a review, a light microscopic review of the specimens that we receive are 20, 30 or 40% tumor content.

And then -- so that's a tremendous determinant of how the allelic fraction might be reported and there's also both light microscopic ways to estimate tumor content and computation ways to estimate tumor content.

My feeling is that in general it is a research tool presently. If -- the last thing we would want is a doc to get a report back that his patient had some oncogenic variant that was unambiguously tied with therapeutic but it was only at 8% so he didn't try uh, you know, a certain TKI that has a very high response rate because that would probably be ill founded.

But I think that's certainly an area in which there is need for data collection and there may be some settings in which that has already been shown to perhaps influence outcome.

The other piece, just tangentially is I don't think that affects, except maybe in the resistance setting, the choice of therapeutic. If mutation A at, you know, 37% and mutation B at 24% -- well then you have to weigh which one is more likely a bona fide driver, which one has a better therapeutic, et cetera.

So a lot of confounding issues there that --

and opportunities that are research questions.

DR. DEEKEN: I would just echo that. I think that's a crucial question that we have no public data -- at least not much on yet. Hopefully with the 6,000 patients tested in the NCI match trial and others that answer will start -- and knowing if patients were assigned to treatment and how they have done that treatment -- there will be some evidence coming out on that.

But I think right now we don't know what to do with that and all the algorithms that are there -- at least in those basket trials, do not incorporate that.

Obviously it's critical that the original pathology investment is as cancer content rich as possible but I think that's a large unknown I think, for the practicing world, from my perspective.

DR. BERGER: Yes, so thanks again for the question I'll add a few things. One is um -- in our own experience we initially were very reluctant to report the mutation allelic fractions because of the concern of how it might be interpreted or misinterpreted because there are a lot of things that determine that level.

I think the low tumor purity or tumor fraction is probably the biggest determinant but there may be tumor heterogeneity -- some mutations may be sub-clonal.

Also, just copy number alterations in the tumor can lead to an increase or decrease in the allele fraction, even at a set purity. So we were concerned that too much would be read into this and there's not -- I would agree, that much data suggesting that patient's with sub-clonal mutations may not respond to therapies that they would have if it were clonal.

Although there are reports, and we have a paper coming out tied to a particular basket trial in the next week or two where we were able to analyze and see a difference between patients with sub-clonal mutations and clonal mutations with respect to response.

So I think this is active research that is going on and we may know more in the future about the effect that that may have on therapies but it might be therapy specific or tumor type specific or mutation specific.

So um, after the concern about sharing this information we actually did get a lot of requests from clinicians for that -- some with a clinical question in mind, some with a more research question in mind so we've begun making that information available to them with some descriptions as to what it should or shouldn't be used for.

But I also want to add that you know, depending on the test the precision in determining that may vary. So for a test that used deep coverage sequencing I think -- MSK's foundation wanted two examples.

You could actually determine the allele fraction with pretty high precision which can help infer the zygosity, the copy number and the clonality of mutations in a tumor which allows us to conduct these research projects and there's some really exciting basic research that's happening at Sloan Kettering with respect to that.

But other tests may -- if they don't have the deep coverage or depending on the nature of the amplification method may not be as precise in driving that value in the first place so there's a risk associated with that where mutational allele fractions may not actually be calculated with much certainty.

DR. BEAVER: Okay thank you. So at this point we could open it up to questions from the audience -- if you'd either go to the first or second microphone and then try to get you in order, yeah.

DR. TSIMBERIDOU: I have a question for Vince. So with the introduction of a high tumor mutational load, there are challenges in selecting the optimal treatment matching molecular abnormalities with therapies.

How, in your opinion, do you have any data outcomes perhaps from your databases how someone should prioritize immunotherapy, targeted therapy? The key issue is that we give these reports to patients and patients are attracted to immunotherapy because these are the normal therapies that can overcome perhaps, a resistance to specific types of therapies also.

So how would you recommend to interpret this data and prioritize treatment?

DR. MILLER: Well unfortunately much of the, you know, the reports we provide from those you know, hundreds of thousands of cases are uni-directional right -- we don't have the clinical follow-up on them.

There was a sub set of cases in part through various registries in which we have participated or will participate our precision medicine exchange consortium, academic collaborations and/or our flat iron health partnership where we have some insights into those.

But I would say the data is -- we don't have a dataset per se. I'd be simply, you know, hand waving based on my oncologic into an experience at this point.

DR. TSIMBERIDOU: Thank you.

DR. BEAVER: Other thoughts about that from the panel?

DR. ROSCOE: I have a question I'd like to ask. This goes back to the allelic fraction. We frequently see kits for sale or companies marketing that they can improve the sensitivity for detecting mutant 100, 500-fold and then we start to get concerned because now you're in a zone where the safety and efficacy of the therapeutic products were never evaluated and even in some cases definite contraindications such as in the case of BRAF wild type, you know where people develop these secondary squamous carcinomas.

And so, actually are you in that "wild-type" zone where this patient can't expect any efficacy but may have adverse events? Can you comment on those types of applications for drastically improving the sensitivity for mutation detection?

MR. BERGER: Sure, I'll start. I think you know, if what's coming from these tests is the identification of a very sub-clonal mutation, and a much lower allele frequency than you would expect based on the purity of the tumor -- other mutations that they're finding -- there's a risk in over interpreting the significance of that, but there may be -- especially if it's associated with emerging acquired resistance, you might expect that to occur at a lower allele fraction.

Having said that I think there are certain applications where this boost to sensitivity is going to be critical like the detection of minimal residual disease, like cell for DNA -- plasmid DNA that the way the tumor fraction is much, much lower than what you typically encounter when you're sequencing tissue.

So the way those technologies typically work -- or at least the ones I'm most familiar with use molecular barcoding and then sequence the DNA sample to a very high fold of replicates.

So each molecule in your initial sample gets a barcode, gets amplified many-fold and then many replicates of each original template get sequenced so that you can eliminate or at least significantly reduce the background sequencing error.

You're not necessarily enhancing your signal but the reason we wouldn't normally calm mutations down to that level when we sequence a tumor is because there's a background error rate for the sequencers themselves that makes it difficult to distinguish false positives from true positives.

So if we can eliminate the false positives produced by the sequencer by sequencing many replicates from each molecule and then collapsing that down onto a consensus sequence that doesn't have any errors, you can calm mutations down to low levels.

So I think it depends on the applications. We need to use those methods for self -- DNA detection, liquid biopsies and for detecting minimal residual disease but maybe not so in solid tumors -- I don't know if anyone would like to add.

DR. BEAVER: Any other questions from the audience, ok?

UNIDENTIFIED SPEAKER: Um, I'd like to go back to something Dr. Deeken said and I think it's very important and that is in this process sometimes you do identify potential germline mutations.

And I'm curious to know what the different institutions are doing to ensure that that information doesn't get lost and that it's being communicated to the patient and their physician that, perhaps, additional testing should be looked at and you know, insuring that there's some follow-through because there is value.

Obviously we want to look at the best treatments available but the ideal goal is to present cancer or to catch it at an earlier stage.

DR. DEEKEN: Just to talk about our experience -- we have our genetic -- our cancer oncologists are part of our tumor program and they'll review all the reports at the time to make sure there's not one of concern that needs further testing, especially additional outside testing for that patient.

And our patients are part of our molecular tumor discussion. The patient and the family can actually be there so they hear that in real time and usually we closely follow that up with coordination with our genetic counselors in terms of further interpretation and testing.

DR. BERGER: So for us our standard analysis um -- masks out germline variants. If the patient is having their tumor sequenced, we sequence the matched normal with the specific intent of eliminating any germline variants from the report.

Nevertheless, every patient signs a consent and that consent specifies whether or not if in the course of analysis, we incidentally find something -- even if unintentional whether they want to find out about that or not.

So um, that was an important piece because we were sequencing the germline, our institution's, our clinical genetic service mandated that.

So for the patient since the beginning who received the standard analysis -- there's no intent to look for germline variants but if we find them there's a process for returning those results back through the clinical genetic service -- through our clinical genetic service.

But more recently as I mentioned we're offering the intentional germline analysis following an additional level of consent. You know initially genetic testing at our center required pre-test counseling and a whole visit with a genetic counselor and medical geneticist and to scale the sequencing that we were doing, that was just completely impractical.

But we actually developed a five minute video that patients watch in their oncology clinic that explains the risks and benefits of the germline analysis.

And after watching that video, the patients have the option of consenting for that second level of germline analysis. And right now I think it's about 30% of patients are opting for that.

And not that patients are turning it down but it's not always offered by the clinician. So it's up to the clinician to decide whether to offer this and it's up to the patient after watching the video whether to accept this germline analysis.

And then we re-analyze the data, report back pathogenic or likely pathogenic variants and that is returned to the treating oncologist and it triggers a follow-up visit with the clinical genetic service.

So it's a nice system that I think has been implemented. It's definitely changed the daily workload of our clinical genetic service. I think in the past many of their visits were discussing hypothetical risks and benefits with germline testing and now, you know, patient after patient is presenting with potentially novel or unanticipated pathogenic germline variant that I think has, you know, really sort of brought our clinical genetic service which in the past has operated a little bit independently more in an integrated way now with the rest of our oncology clinics.

DR. BEAVER: Any comment from Foundation Medicine?

DR. MILLER: So many of you may know that we do somatic -- we do testing on tumor-based tissue and not a matched normal so we're looking for oncogenic variants, you know, regardless of whether they might be germline or somatic in origin.

We do caveat reports when we believe a variant has some chance of being of germline in origin. This becomes of course when one looks at liquid biopsy products a different kettle of fish so to speak, but it's in the ability to think something is germline is far more apparent.

And of course, one of the challenges different from working in a, you know, a single academic institution doing a great test is the challenge of doing matched normal at scale in clinically relevant timeframe, so that's one of the distinctions.

And the converse is the challenge and it sounds like theirs is, you know, working towards a solution of it is in theory -- and I've been in clinical and have seen patients at Memorial who have a -- who've not signed the consent to learn their germline status.

They could have an oncogenic variant that might be therapeutically targetable say with a parp inhibitor and it might be unbeknownst to the clinician or the patient.

So it's a -- it's a challenging issue on both sides but I think this is inherent with a lot of new technologies that come forth. We all used to see CAT scans come back with, you know, suspect PE or suspect coronary artery disease or something where the, you know, the radiologist was you know, thinking outside maybe his area of expertise in some cases or his focus.

And what -- how does one properly address those? What's too much, what's too little and what's right and how do we evolve that over time?

DR. BEAVER: Okay, thanks, other questions from the audience? Um, I'll ask one maybe last question if we don't get questions from the audience we can be a little ahead of time.

In all of the talks we touched on sort of expanding the panel -- keep broadening the panel.

What are some of the other pros and cons perhaps about a disease-focused panel, a smaller focused panel versus these large panels and at what point do you stop expanding the panel or the variants? Do you want to start Mike?

DR. BERGER: Yeah I'll start. I mean I think there's been a clinical benefit to offering all patients one panel because it has allowed us to discover variants in genes that might not have historically been associated with that tumor type but when found in that 1% or less than 1% of patients with a given tumor type would qualify them for a therapy or clinical trial.

So from a clinical standpoint I think there's been a benefit to a large panel. From a research standpoint especially as we're trying to mine the data the fact that -- as Dr. Miller alluded to, patients are sequenced with a uniform platform over time, over an entire cohort makes it much easier to draw inferences from and interpret the research and clinical findings from that cohort.

I think with respect to tumor type specific panels also it um -- it's very challenging with the workflow for the laboratory, especially high through put labs like ours within a single batch of samples in order to, you know, run a batch every day and make sure that the turnaround time is as short as possible, they're all batched together with a single panel.

So I think workflow and operational considerations have driven us towards a tumor type diagnostic thing. Now as we're thinking about cell free DNA and more sensitive assays -- liquid biopsy applications, they are where you need to sequence a much higher depth, you might not be able to afford to be as broad.

So that's where I think tumor type specific panels may re-enter our consideration but still I think the logistical and workflow challenges may prevent us from going in that direction.

DR. BEAVER: Okay, Dr. Miller or Dr. Deeken?

DR. DEEKEN: I just want to point out I think that there's -- by restricting the list too much, I think we lose the opportunity for discovery and for methods of resistance.

If you're only targeting especially upstream key mutations and you're not doing the larger panel you might miss in terms of drug therapy efficacy, because for these targeted therapies they're working maybe in a third of patients, 20% of patients -- some end track inhibitors are doing better than that but we might be missing the discover opportunity by narrowing it so much that we miss the identification of mutations that might be driving resistance or acquired resistance along the way.

So my bias, especially again as the cost of sequencing changes is to not have too narrow of a panel to miss an opportunity -- because if you think about how medical oncology has changed -- we're now doing disease focus phase 1 trials as well as 2's and 3's.

We're narrowing patient treatment so much that we're going to lose the opportunity that often times drug discovery leads to in terms of the serendipity of discovery along the way in terms of what might work in a disease you didn't think about or a mutation you didn't think about.

So the narrower the list I think the reduced opportunity, especially in clinical trial settings of making that additional understanding of pathways and activation and what might warrant resistance with targeted therapy based on just a narrow panel.

DR. MILLER: So I agree with what both gents have said and they certainly speak to our approach. I would say the greatest challenge or push back we get is often around turnaround time and even though our turnaround time once a sample is received maybe quite clinically relevant -- say 10, 12 calendar days.

One doesn't know what happened beforehand. So certainly as testing moves as part of the workup, pre-frontline therapy and metastatic disease that takes away some of this challenge although we continually push to shorten our turnaround time.

If you have a patient with advanced cancer who you are only thinking of doing testing on when they have failed their first line chemo and they have a crescendo of symptoms, it's sort of game over because you're unlikely -- it's like an eclipse.

You're unlikely to have a phase 1 trial matched to that patient's tumor open at your institution that he or she can start in the two weeks before their symptoms go from bothersome to, you know, to declining performance status, clinical trial ineligible.

So make it a chess game and a strategic decision to test up front, whatever assay you're using and don't do it after, you know, the individual is basically in extremis.

DR. BEAVER: Great, we have a question from the audience.

DR. LICHTENFELD: Thank you, Len Lichtenfeld, American Cancer Society. I appreciate the panel and I suspect this is a theme throughout the entire day.

One statement, Dr. Miller you mentioned about the need for both regulatory and payment processes to modernize for lack of a better word to have the resources devoted -- to be able to respond to the rapidly changing signs -- a critical issue not only here but with a number of other arenas within cancer care.

But the question which you may or may not wish to respond to at this point is that you know, what we're hearing here from some outstanding institutions and companies in terms of what you do and how you do it and how you curate and how you validate and what you say, what you don't say.

However, you're not alone. There's a big area -- a big industry out there that's trying to guide people to what kind of treatment they receive, some of whom are doing their own variant analysis and saying that we think this is something that you need to pay attention to.

And certainly consumers -- patients of course, and the clinicians who care for them, are not really as up to speed. What do we need to do to make sure that everyone -- and like I said this may be the question of the day -- what are your thoughts about what we need to do in order to ascertain that the information provided is truly clinically relevant and actionable in a genuine way?

I'm thinking about -- Dr. Deeken what you said about Inova. But you know a lot of care -- Inova, I don't even consider it a community hospital system anymore -- it's really a major institution. But there are a lot of places out there in a lot of parts of America that just don't have that expertise or capability.

What do we say to them and how do we make sure that the care they receive -- the information they receive is in fact, accurate and actionable if they have the resources they need to provide care to their patients, thank you?

DR. BEAVER: Who wants to take that one?

DR. MILLER: I think I'll recuse myself from this one but those in the room I think have outlined a great path to both for academics and for profits to create a, you know, a high bar but a doable one that ties together payment with knowing what one is doing or not doing and is providing to doctors and patients.

DR. DEEKEN: I mean I think that's a key question and it's not the wild west out there but there are companies popping up all the time and I think that's why regulation and evidence and the leadership of our major cancer centers in this country need to help set that tone and from a regulatory standpoint why it's so critical.

I remember being a, you know, new oncologist in 2006 and a patient brought in a Caris report who the surgeon had ordered -- and that was their creative approach, you know, to start this field to not go to the medical oncologist but to go to the surgeon.

Now of course we were all poo-pooing it then and when we started working and opening up this whole new field we changed our tune. But I think we need to be aware of -- and the later session today I think is so, so critical in terms of the knowledge based and levels of evidence because you know, we're getting that point that we got to in terms of evidence levels that should drive treatment decisions and what's an unknown significant versus a known significant variant.

That's where the science and the -- you know, the translational science, basic science clinicians have to be so critical to help us define those and have agreement on those so that companies aren't making it up along the way.

Because knowledge based curation is incredibly time intensive if you do it right. And if you don't do it right then you're going to find recommendations that don't fit potentially.

DR. BERGER: Yeah so I guess I would echo that final point just that -- and I don't have the answer and I think this is a big question throughout the day.

It's just very hard and it's a lot of work, a lot of time, a lot of expertise to create and a lot of resources to create a knowledge base and it would be nice if everyone didn't have to do it and then I hope that not everybody has to do it and I think in our own institution we deliberated a lot

-- was it going to be worth the investment.

But what happens is you see reports issued from other academic labs or other companies and you disagree with the curation and I think that's what prompted our center to invest in developing and curating the expertise throughout our clinicians and it's not all correct and it's not comprehensive.

So you know, I think probably a theme of the future -- of the next few sessions will be how to leverage those efforts across different centers, maybe share this knowledge base, come up with community standards for that.

I think that's where we ultimately are going to need to go.

DR. BEAVER: And Donna, did you have a comment?

DR. ROSCOE: I was just going to chime in with a plug for the discussion later on today and say that that's what the agency in conjunction with a number of other agencies and experts are working on creating a database that will be accessible universally to everyone.

DR. BEAVER: Great and from the audience?

UNIDENTIFIED SPEAKER: So I have a question from the perspective of someone who -- sure, hi, so I have a question from a perspective of someone who would ask the deep dive for knowledge, curation, evidence -- that determination.

And I think one of the fundamental differences between germline and variant -- somatic variant classification is the literature itself because most of the germline curation happens at the -- there's a lot of evidence at the variant level because it's variant transmissions, functional transmissions, functional meiosis, you know, controlled and you know, functional studies at the variant level.

Somatic variant classifications are you know if you notice the literature at the individual variant level may be very sparse -- it would be more at the gene level or the domain level at the exon level or even broader and mostly the variant effects are never -- at the variant level they're qualified by the complex molecular profiles of that variant in variants in patients with rearrangement.

So it makes it a little difficult to compare studies and, and really piece apart the variant contribution to action-ability, you know, sort of agnostic to varying molecular profiled reported in literature.

And there is also, you know, the need for standardizing what constitutes a small clinical study, what constitutes a large clinical study, especially when a patient with a variant in that large clinical study maybe just one patient, but the study may be very large.

But yet you're classifying a variant. So one of the approaches in terms of weighing the somatic -- the diversity of somatic evidence in the literature in, in coming up with meaningful classifications which was a rather broad question but I hope --

DR. BERGER: I'll give it a fresh shot, thanks for the question. Um, I think um, you know the most comprehensive of these knowledge bases for somatic mutations are attempting to annotate at the variant level.

I think you have to acknowledge -- you have to annotate at the variant level because you can have passenger mutations, driver mutations, activating and inactivating mutations in the same gene.

So part of the OncoKB curation involves variants that may have been functionally characterized either in vivo or in vitro.

Having said that I think especially as we accumulate more data -- 20,000 cases, 100,000 cases, 180,000 cases we can use statistical analyses to identify hotspots of mutations.

Mutation is -- we currently observe more than you expect by chance and we have good methods for doing that. And we have been identifying novel hotspots from larger analyses that are found.

And when that mutation is found in a patient, there are instances where patients have been enrolled on a trial or received a therapy at Sloan Kettering without any evidence of in vivo, in vitro of that specific mutation conferring sensitivity to the drug and they've responded.

So I think, you know, part of the new paradigm may be to use the large scale, you know, big data approach to identify hotspots and then directly test in patients.

That has worked. It doesn't always work, it might not be a universal approach but I think we can leverage the large-scale sequencing data that's being generated provided that it's being shared through the genomic data comments and AACR GENIE and so on to try to infer function and at best, or at least prioritize which mutations may then go on to functional characterization for the definitive evidence that may be published that you may be looking for.

DR. MILLER: I would just echo I agree with Mike the complexity of somatic variants mandates in part that a lot of our learning will be going back from clinical experience and therefore universal reporting -- reporting variants in the same way which is you know, will have been addressed in other settings.

It is essential real world evidence, has been an FDA initiative and in particularly an oncology to consider as drugs are getting approved appropriately for, you know, on fewer and fewer patients when there’s a you know, clear, robust signal of activity.

And of course the other piece is germline testing and of course it’s incredibly nuance and this may be an oversimplification but some level is more simple than the diversity of what we see in somatic alterations and cancer.

And the field has also been around doing this routinely on large numbers of patients for many more years so there’s also a sort of out there first piece to this.

DR. BEAVER: Great, thanks so much everyone. I think we’re at our time now so we’ll have a break until 10:25 and let’s give our panelists a big round of applause.

(Break)

DR. MADISON: Just one housekeeping thing. If you come up for the question and answer, to help those who are online, could you please just state your name and your affiliation before you proceed with the question so we'll have that within our transcript.

And we'd like to welcome Dr. Anand Pathak, he's the Medical Officer in the Center for Devices and Radiological Health in the Division of Molecular Genetics and Pathology. He'll be the moderator for Session 2.

DR. PATHAK: Hello, good morning. It's a pleasure to be moderating this session, session 2 - Levels of evidence required for reporting variants and guiding patient treatment.

So as mentioned during the Q and A session of the last talk, the information about whether a variant is truly actionable or not and whether or not that information is valid is critical to patient care.

And to address the issues of the rules of evidence we have five distinguished speakers. Three associate speakers and two panelists from across the country engaged in clinical practice to provide their perspective of how they use levels of evidence.

Our first speaker is Dr. Sahikant Kulkarni. Dr. Kulkarni is a board certified medical geneticist trained in clinical molecular genetics and clinical cytogenetics. He serves as a Professor and Vice Chairman in the Department of Molecular and Human Genetics at Baylor College of Medicine.

He's also Chief Scientific Officer at the CAP CLIA Lab at Baylor Genetics, so please welcome Dr. Kulkarni.

DR. KULKARNI: Thank you Dr. Pathak. It's a pleasure to come here and share our experience at Baylor College of Medicine. So what I was asked to do today was to give a workflow overview of our unique um, organization which is a hybrid organization which is organic as well as for-profit commercial, clinical, diagnostic lab. And then I will share some of the early work we are doing with ClinGen and ClinVar in somatic relation related to the curation and making harmonized standards for curation.

These are my disclosures. So just before I start that part I will give you a brief -- very brief, overview of our lab. So we are an organization which does full service clinical genomics testing for all stages of human life and for using all the different tools which are the most cost effective tools starting from pre-conception to cancer.

We are a relatively large organization. We have about 300 employees and 25 directors who are board certified and molecular pathologists. And we are extremely lucky and we have derived a lot of benefits from a very strong academic center. So our department -- molecular human genetics has 180 primary faculty members. And the symbioses between the basic research, clinical genomics laboratory and clinical genetics testing and about 35 genetic counselors and a genetic counseling program -- I think we have a very good understanding of doing these kinds of testing.

So we have -- when I think of testing um, we were the first ones to launch a non-invasive sequencing testing in pre-natal called PreSeek and of course we do cancer testing.

So I was asked to give one clinical case as an example to kick start the discussion and to wet the juices if you would. So I wanted to start with the case which is more of a research case but exemplifies my dream and my ideal situation going forward.

So this is Dr. Lucas Wartman, he's an oncologist himself -- he's a leukemia doctor. His story was published in the New York Times about 3-4 years ago.

So he developed with pre B-ALL, acute lymphoblastic leukemia. His cytogenetics was not pathognomonic but all he had was one deletion of the short arm of chromosome 12 which took the gene called edv 6 but it did not rearrange it, had several lines of therapy and so therapy in 2011 had a CNS involvement after um, sibling transplantation and at that time it was decided to do whole genome sequencing and RNA-seq in this individual.

So there were a lot of different sequence alterations, fusions which were found but nothing which was clinically actionable. But one thing which was found only by doing RNA sequencing was FLT3 overexpression and it was known -- this alteration was known to be sensitive to FLT3 inhibitor Sunitinib which is of the up-root in renal tumors but not in this particular tumor.

So um, long story short, Sunitinib was given to him and he responded very well back at work writing grants, seeing patients. So this is an example of how it should work where you do a ban genomic analysis and have that ability to not only find new treatments but also use this information for disease monitoring.

So now there are lots of ways using variant infrequency, number of blast counts, variant frequency in overall genome and also using fish based assays to detect the response of the therapy for individual deletion.

So this is an example which exemplifies where it should be heading towards where you not only use the genomic analysis but you have the ability to constantly monitor using different approaches.

So this leads me to now segue into the clinical challenges and what the challenges we face in the clinical diagnostic lab. But I'm not going to talk about all the challenges but I am going to focus towards the end on the lack of standard and guidelines and what are we doing as a society to help that.

Just if you have seen these kinds of slides before our workflow is very similar to a lot of other previous speakers. One thing which we decided to do differently in order to get a better turnaround time is to use a field programmable gate array based approaches.

Essentially it streamlines the workflow and the hardware and the software is encoded in this chip and this is one of the first few applications in genomics FPGA is widely used in other industries like um, aeronautics and space exploration.

So our approach is a UMI based approaches which has UMI's on both the ends. This helps in detecting very low level variants and it has been in our hands shown to detect very, very low variants and I think this would be a very good approach for our liquid biopsy tests which we are going to launch soon.

So we have about 277 genes listed here and this was done in a very extreme vetting way. We asked our colleagues -- clinical colleagues, oncologists, molecular pathologist, both at NCI designated cancer level -- center level, also at community level to see what their wish list and then we did an extreme analysis on the clinical validity and utility of those genes.

This is a screenshot of our report. We also had um, a need to make it very, very concise and this report before it was finalized was sent over to about 60 different cancer centers and community hospitals all over the country to get feedback.

And so we have -- the first page is very simple with the main summary and a brief paragraph of the interpretation and then we have detailed analysis and the AMP classification based analysis and evidence based details in those subsequent paper's pages.

And then of course, it also talks about the clinical trials. And here we have done -- and I'll show you in the next few slides we have gone a step forward and since we have a reference lab we deal with other hospitals.

We have worked with them to establish API's where we can directly feed into local and institution specific clinical trials in here.

So I wanted to spend a couple of slides to talk about what we are trying to do in putting it all together with the help of a program which we are launching with two major healthcare systems to democratize the access but as one of the audience questions was how to make it more community wide and how to democratize access by making sure that the correct information is relayed and nothing is lost in translation.

So just to give you the scope of the project we have 70,000 new analytic cancer cases in this system and the total number of lives affected is 95 million, so it's an honest opportunity to impact and make this precision medicine approach available at a much larger scale.

So this is done through a conglomeration of different API's including the lab, the epic based or any EMR based systems, EMR's, cancer registries, clinical trial management systems and it's all put together to have more detail and searchable reports which have details shown here.

You have the ECOG scores, the size of the tumor, date of the test, all the different drugs, different tumor boats which were a part of it and this is all recorded in the system and is -- is available for the oncologist and it's available for the lab as well to look at all this data back and forth.

You can -- the clinician oncologist has the ability to record notes and request a tumor board specific for that patient which is more like a consult but has the ability to invite other professionals as part of the healthcare system to collaborate and make that part of the patient record.

And if an oncologist is interested in um, searching the whole healthcare system on patients like mine they have the ability to do that and they can see um, patients with similar molecular alterations, similar tumor type and also look at the outcome based on not only that particular treatment but based on other treatments which are part of the -- that particular patient's.

So I think now is the time for me to talk about the efforts which we are doing as a community. Many of you might know and I believe Heidi Rehm is representing ClinGen and was going to be talking about this in detail but what we have done is taken that part of the -- one of the clinical domain scope ClinGen Somatic Workgroup and there are several people in the audience who are part of it and are actively contributing to this and we have been active for 2-3 years now.

We have members from industry, academy and government and so the aim is to have the annotation and interpretation standardization in cancer somatic variants and to come up with more detailed harmonization with different guidelines.

And we have monthly calls, we have face-to-face meetings, we have met at different meetings like AMP, AACR and we have another meeting coming up in next year -- this year, AACR in April.

We have done a lot of work up until now. We have developed a minimum variant level database which we can see and we have a lot of different task forces -- oops, a lot of different task forces which are disease specific.

So we are working to create this whole ecosystem where we can put it all together. These are the 18 data elements which we are standardizing awaiting submission and we use the AMP guidelines and we are modifying that as we go.

So this is the paper that was published in Genome Medicine. And then I just wanted to mention that AMP led an effort which had multiple institutions and we have talked about this morning to harmonize and standardize the different tiers of variants and so it started with a survey -- a membership-wide survey on the ways the variants are reported.

And as you can see here, there were a lot of -- I don't have time to go into details but there are a lot of variability. So we came as a community together -- as I said there are a lot of people in this room who participated in that and we came with a variant level classification -- evidence based classification of these variants in four levels.

And I don't have time to go through this -- this was published, but tier 1 is FDA approved level A and then tier B or tier 1 level B is the other evidence from our studies and so forth. So this is all published in the *Journal of Molecular Diagnostics* in January of 2017 so you're welcome to look into it.

And we continued as a ClinGen somatic working group we continue to finding the ways we can put it all together and these are the acknowledgements.

I'd like to just say one thing here -- one of the major problems of curating these annotations of course is the lack of funding and NGI has been very kind to support the ClinGen group but we are also looking for funding from NCI.

And the other big bottleneck is the lack of incentives for the submitters to submit the variants within the ClinVar database and so we're working with a journal called *Cancer Genetics*, um, I happen to be the editor-in-chief and so we have created a system where the submitters can submit a variant in a very standardized format using NPTS system and the NVLD and then we are hoping to build API which will help the variant submission directly into the ClinVar.

So the net result is that we get a good variant clinical evidence driven variant information with outcome and methodology and flow, whatever clinical information we can get but at the same time gives the department ID for the investigator who is submitting the variants that has a win/win situation. I'll stop there, thank you.

DR. PATHAK: Yes thank you Dr. Kulkarni. Our next speaker is Dr. Howard McLeod. He's the Medical Director at the DeBartolo Family Personalized Medicine Institute.

He's the Chair of the Department of Individualize Cancer Medicine. He's also a senior member of the Division of Population Sciences at the Moffitt Cancer Center, thank you.

DR. MCLEOD: Thank you, it's a pleasure to be here and I'm really glad that this topic is being addressed. It's something that is not going to be easily achieved because there are multiple different groups of needs that are coming out of this.

So we've heard about and talking about the needs of the laboratory in terms of producing a well annotated genome that can go forward. There's also the clinical need -- not just in terms of picking what drug, but in the context of all the other things that are happening with the patient.

And so there will be aspects of this that can be codified, put into guidance, put into whatever else needs to be and there will be other parts that will need to be the practice of medicine because that patient's renal function is different than someone else's or whatever it might be.

And so I think that will come out shortly -- a little bit in this presentation and I bet in the next one as well.

Also within the clinical problem is there are multiple active regiments for the treatment of most diseases and so it is rarely a choice of good drug versus no drug or even a good drug versus bad drug.

But it's almost often -- almost always a choice among equals -- two really good options and you have got to pick one. And typically we'll pick the one that we know how to spell or we'll pick the one we're most familiar with or we'll pick something based on less than objective data.

And with genomics certainly we're edging towards objective data but that's always still part of the problem.

Also, there's great variation. Even our best examples -- our homeruns are 80 or 90% successful -- most of the time we're 30% successful -- variation does exist and so that has to be factored in -- that we're not looking at perfect, we're looking at good, on the way to great, someday perfect in our decision-making.

We also cannot ignore toxicity -- I'll hit this in a moment, but we talk about risk benefit ratios and then we only talk about benefit -- we don't really gauge the risk part.

So we need to be bringing in the patient not just the tumor in terms of the discussions. And of course there's a part that no one wants to talk about and that is these therapies -- especially some of the newer therapies -- even if you're well insured your co-pay may be $2,000 a month and that is something that most Americans cannot come up with lightly and therefore we need to be having rounded decisions about the patient and what the burden is -- not just looking at these things in isolation.

So selecting from amongst equals -- our study designs are not made for that. If you look in a copy of *The New England Journal* it will come out Wednesday night -- you will look at it and there will be a winner and a loser from that clinical trial.

There will be a kill curve a Kaplan-Meier curve. It will show the winner and the loser and that will be the punchline from the story -- in reality that's first line and second line therapy.

You don't go to a patient whose first line therapy stopped working and say, "Sorry we're going to try loser therapy now." And so we create these models whereby it's a winner/loser study design and then we go out and apply them in a very different clinical situation and then we're wondering why we don't have the ammunition we need to help patients make a clear decision.

And so we have some work to do on the study design aspects, not just in terms of the way we develop drugs. We already talked about it in the previous session -- the anatomy versus non-anatomy based approvals and that's an opportunity but again will not be the norm in my personal opinion -- and then the toxicity part is there.

And this is just a reminder that our patients have at least two genomes which we need to care about. They're tumor genome which is probably multiple genomes and they're normal genome and of course there's the microbiome and a bunch of other genomes that are important as well.

And typically we try because we want life to be simple -- just focus on one little aspect of one of these things. Really we need to be tackling this in a further -- in a broader way and it's just so hard that we don't want to. Even drawing this figure was hard because there's such a divide between the germline and the somatic people.

A somatic genomics person does not want their daughter marrying a germline genetics person and vice versa. It’s a religious divide but the patient has both and we need to tackle that and that’s the opportunity.

Also with the FDA approvals, this is looking at the dosing and administration section -- it’s about 160 drugs with genetics somewhere in the package insert - this is in the dosing and administration section.

Many different examples involving the tumor but of course also for cancer patients, a lot of the other stuff that we care about in terms of managing these folks have examples in there.

And so there’s an opportunity to really think more holistically about how we take this forward. Also there are a lot of different patients -- we’re not Memorial Sloan Kettering, we only have 120 patients a week -- not 150 to 200. Some day we aspire Mike -- but within that there’s a subset of those that get a more intensive review and of which is about 30 or 40 a week that really drill in with the personalized medicine team, but it’s a growing number of patients.

It's not something that is going to happen some day or has plateaued out in terms of the opportunity. It's also lots of different types of tumors.

This is just showing the incidence of testing in the last year and the different types of tumors there -- um, a lot of different kinds of cancers getting tested, not just all one particular type of cancer.

Now the way it's actionable in our institution -- at least the way I'm focusing this talk is around a couple of different ways we're using NGS in that way.

One is for a benefit or resistance to a particular therapy -- think KRAS and (inaudible). FDA approved therapies -- think non-small cell lung cancer -- we have to look for lots of different options, but clinical trials and that's certainly an important part and we're going to hear a lot about that in the next presentation -- a very critical part there.

We too are somewhere in the 11 to 15 -- 10 to

15% range in terms of patients going on trial. And then the use of FDA approved therapies for off-label types of cancer -- and I'm not sure if we are supposed to talk about that within this building but the reality is there are a lot more patients that fit that criteria than they do clinical trials because it's a third cancer or they have bad kidneys or brain med's or some reason why they can't go on a clinical trial even though their molecular status makes them eligible.

There is also a lot of prognostic information within the HEM side of our work. Often we're trying to figure out someone who needs to go straight to transplant rather than get chemotherapy or biological therapy -- that's an important therapeutic decision as well -- it's not just all a choice between drug A and drug B.

And then lastly the germline aspect can be quite important and we've hit on that already. We look at two different levels of evidence -- now I should say within the laboratory setting we use the AMP CAP -- whoever else was involved guidelines.

Our's is the AMP part but I know there are others involved, and in terms of trying to annotate a variant. But then there's all the other aspects that are there.

In terms of what sort of supportive care data we need to then think about clinical action-ability. An important part of this decision with our institution is having access to the clinical record and understanding what they've already received, what their organ function is like, all these other features.

And I mentioned that because it's going to be very hard in the context of guidance to put all those features in. Secondly, if you're a testing company, you may not even want that data because of the liability it brings and so they're aspects of that that we're going to have to be preparing for a local site to take advantage of -- even though the testing company or the algorithms may not exactly encompass all that information.

Also, our clinicians need to make a decision on the patient. The patient is fit, they want a therapy, they're going to pick something and so they're looking for a feather to tip the scale.

Every once in a while we'll find an amazing therapy that is the answer and you have to give that. But back to this choice amongst equals -- if you have two equal options, a feather will tip that scale -- it does not need to be a two-ton weight.

And so the type of data, the weight of impact can be as little as case studies, case series, or even in some context pre-clinical data to break the tie.

It's not what we're looking for -- it's not our goal. Our goal is level 1 evidence but when it comes down to it, that oncologist is going to make a choice. You know the old Rush song -- "If You Choose Not To Decide You Still Have Made a Choice."

We are going to make a choice. So can we use the data available in terms of those levels and so we've designed an approach whereby when the testing comes back we have a personalized medicine consult service or clinical service that reviews every case, post-laboratory.

The laboratories are involved in the discussion but it's more a therapeutics review than it is a laboratory review. For a special case, it will be pitched out -- well I'll come back to that to our version of a tumor board -- but then it goes through the process.

And so we use these types of level of evidence -- we have 9 levels which range from FDA approved drug for that specific cancer -- thank you, all the way down to no information is available.

It's important to say that as much as it is when there is data available because a decision, as I mentioned -- as I belabored, a decision will be made. Um, and so what the data is there.

And so that allows us to put together recommendations as shown here where there might be multiple different therapies at different levels that can be recommended and we'd also take on the germline piece as the small little -- the second paragraph hits on -- in terms of making sure we don't ignore that part of it.

We've all found tp53 yolefame mutations on page 7 of some commercial lab's report because they didn't want to face the fact that there was something toxic there and tried to just ignore it and so it's an important thing -- we have to be taking into account.

Every week all these cases get reviewed in what would normally be considered a tumor board, but then for the special cases that need a deeper level of review, we have something called the Clinical Genomics Action Committee.

Now, it's a supermarket tumor board but too often molecular tumor boards -- especially academic centers, are really just freak shows. We're looking and saying, "Wow, that tumor has a JAK2 amplification, isn't that crazy? JAK2 -- next," as opposed to what do we do for this woman and how do we treat her cancer.

And so, by changing the name we've also

changed the mindset of the focus and belabored that point and lots of different disciplines involved trying to make these decisions.

We serve drinks and cookies that's why everyone is smiling but the idea that these different people who may not discuss cases together can weigh in because a variant seen in a leukemia patient may suddenly appear in a sarcoma patient and we need that cross representation in order to try to really interpret how to go forward and it's put into the package.

This is my last slide -- or last thank you's. So just a reminder it's really a choice for amongst equals that we're there. Clinical trial options are paramount, how do we make better decisions? The longitudinal monitoring for futility or next options is really important.

CT scan is yesterday's technology. Can we use some of these molecular approaches to better make these decisions and then of course toxicity is something that we do a great job of estimating ourselves and a lousy job of estimating from the patient themselves and so there's an opportunity there.

So I'll stop at that point. There's a thank you -- I thought I put it in the thank you slot, there's a whole bunch of people involved and I'll go on to the next presentation, thank you.

DR. PATHAK: Yeah, thank you for that presentation Dr. McLeod. Our next speaker is Dr. Tsimberidou, she's a Hematologist/Oncologist and Professor in the Department of Investigational Cancer Therapeutics at MD Anderson Cancer Center.

She also initiated the precision medicine program there in 2007, thank you.

DR. TSIMBERIDOU: Thank you for the invitation. Today I will share my experience starting with the program of personalized medicine and the challenges we have using levels of evidence required for reporting variants, how we guide patient treatment and I will answer also certain questions for this presentation.

So the first question was how do we incorporate levels of evidence into variant reporting and clinical practice? The question is how do we define precision medicine because in 2011 for instance, the definition of NCI included a form of medicine that includes information of a patient's gene's protein's environment to prevent, diagnose and treat cancer.

In practice to implement personalized medicine we need to have a tumor molecular abnormality that does not inhibit -- that is known to cause cancer and also we need to have drugs that are successfully and effectively in keeping the function of the genetic alterations or the biologic abnormalities.

We have to be able to use these drugs consistently and effectively. The current definition after the introduction of immune oncology drugs in recent years includes the use of therapeutic agents that target any biologic abnormality that's associated with carcinogenesis including immunotherapy.

How do we use different levels of evidence at

MD Anderson? We order molecular profile as a standard of care using our, for instance, internal profile or other molecular profiling available.

Many patients are referred to us or they come with multiple -- sometimes molecular profiles performed either in their tumor or even in their cell free DNA analysis and we need to take these data into consideration when we determine how to treat them.

Also we have molecular profiling done as part of clinical trials for instance the IMPACT2 trial in the center for molecular profiling and advanced cancer therapy that I am conducting with sponsored in part by Foundation Medicine -- we have molecular profiling, the NextGen sequencing profile and some also markers like tumor molecular mutational load, MSI status, PD1 - PDL1 status, also the NCI-MATCH MPACT trial, attract molecular profiles that are done as part of these trials.

How to interpret the molecular profile -- we take into consideration our expertise -- or based on the expertise of oncologist, precision oncologist and in our clinics and also we have a specialized team of molecular biologists that interpret the data.

How do we select the treatment of patients treated on clinical trials? First we take into consideration the recommendations of tumor molecular boards.

We have a tumor molecular board, for instance, in IMPACT2 we had the board every two weeks and then on a weekly basis we discussed in our department of investigational cancer therapeutics, the specific molecular profiles, the annotations, all of these abnormalities and also we select -- we propose treatments based on their variable clinical trials.

We have to have -- then we screen these patients who have to have clinical trials available. We discuss with our patients -- we screen them so it is based also -- the selection of treatment on trial's availability and patient preference and we have to have sponsor approval as well as more importantly insurance approval to enroll and treat patients on clinical trials.

We have a stringent regulatory CRC/IRB/DSMB review of our clinical trials. Our review is included in all randomized trials and also we have trial prioritization that depends on the department, the timing and other factors that are internal at MD Anderson.

We have a specific -- a precision oncology decision support team that helps with the annotation of the clinical abnormalities and um, this support team aggregates data directly from clinicaltrials.gov and also if the trial was not active for two years at least the status is unknown, then it extracts the state names -- for instance if it is looking for trials that are conducted in Texas and the users sort data from various databases as you can see here -- COSMIC, National Library of Medicine, the European Bioinformatics Institute, ClinVar, dbSNP, Ensembl and the National Human Genome Research Institute.

So then once these reports are generated and a report is sent to the physicians and the physician discusses these annotations with their patients and doctors they are discussed as mentioned in weekly conferences, all the disciplinary conferences in order to prioritize -- not only select treatments but also prioritize them based on information provided about business status morbidities and brought up to its ability.

How do we implement the CAP, AMP, and ASCO recommendations? So um, we -- probably about a year ago the different variants and how we prioritize them, we use in general these data. However we also take into consideration the continued involvement of data indicating that we can -- that a gene for instance, that's in alteration should be targetable and of course the term targetable includes available clinical trials with drugs that are known to inhibit the function of the gene.

And the challenging question is what -- how do we design clinical trials or how do we use the level of evidence to include basis of clinical trials?

In recent years in addition to tumor molecular profiles there are protocols who allow patient enrollment based on their cell free DNA analysis and as discussed earlier we have at times um, altering profiles from both tumor and cell free DNA and there is a lot to learn and we continue conducting clinical trials to understand the clinical significance of these alterations.

The old paradigm from moving to tumor type has changed completely and therefore now we deal with multiple alterations, this is a simplified schema slide of patients with lung cancer and there are mutations we see in several of these patients.

Also, more importantly with the introduction of immunotherapy PD-L1 and other new markers overlap with several of these alterations.

This is an example of a patient of mine with salivary cancer, treat with a BRAF V600E mutation, for instance, who was treated had um -- was treated with Vemurafenib but on a basket trial and had a complete PET response after two cycles of treatment -- so we have seen the model of changing from tumor, treating the tumor type to treating tumor type with a specific molecular alteration and seeing successful results in basket trials that we would not have seen otherwise if we did not have drugs available in this trial setting.

The last and most important perhaps question that I was asked to address in this presentation is what can the FDA do to move the field forward?

So the current status is that we have multiple clinical trials with correlative scientific endpoints, that there are dynamic changes in time and space of tumor, microenvironment as well as the circulating tumor CAN.

And in addition there are complex molecular networks, immune mechanisms, proteomic, transcriptome and epigenetic changes and we identify now multiple tumor and ct-DNA alterations and as well as immune abnormalities in individual patients.

There are prospective clinical trials with adaptive design that we hope they will accelerate the drug approval process by reducing cost, time and number of patients.

So what can the FDA do about this field? We believe that we should facilitate the approval of platform diagnostics rather than requiring drug-specific companion diagnostic.

We should make the tumor NGS available to all patients and we're very pleased to say recently that some time which to be approved however this does not cover all bases with any tumor type early at the stage of the disease and it is definitely not covered by patient's insurance for all patients which would accelerate drug approval across tumor types based on biomarkers.

And we should continue to raise awareness to drug development ECO-system for the most efficient methodologies to determine effectiveness of novel drugs.

Those that are involved in drug development including investigators from pharmaceutical companies should be in discussions with the FDA to know ahead of time, even before they started designing a phase 1 trial, what is required and what is considered effective in order to design and conduct an effective clinical trial that will end to the FDA approval of the drug or drugs.

Also we should encourage basket trials and combination regiments with innovative design to expedite biomarker-based drug development rather than testing one drug for one marker for instance.

The FDA should lead the evolution to transition to the new environment and pharmaceutical companies -- although the FDA has shown a lot of flexibility and we saw a lot of approvals recently -- in pharmaceutical companies there are still a lot of resistance to move forward and many of these companies including statisticians and other um -- other people who play a major role in protocol approval, they are not flexible with protocol designs.

The FDA should continue to work and education and lead this transition to the new environment. Also they should provide leadership by encouraging alignment in philosophy between FDA and IRBs.

Often we have noticed delays in the IRB approval of protocols that cause non-value enhancing delays for protocols that are already FDA approved.

The FDA should require minimal, essential data to decrease cost and complexity of trials. We're all aware of how the data that are not necessary can drive the cost and have, for instance, major CRO's that increase significantly the cost and the time required to complete the trial

And to encourage sophisticated phase 1-2 trials with innovative design. They should help develop innovative bio-analytical methods that are better adapted to the current precision medicine environment for new classes of drugs such as for immune oncology, trials with new endpoints such as a two-year landmark analyses may be more meaningful than just a plain log-rank analyses.

This would continue to utilize the rapid approval process such as the breakthrough designation and fast track programs. For rare mutations, the FDA should consider moving towards increasing use of expanded access patient data contributing to the approval of precision drugs for rare molecular alterations and diseases, rather than only considering patients treated on clinical trials.

And the expanded access program should be simplified. Also the FDA should consider developing a novel pathway to expedite drug approval based on successful results of well created and "N of 1" databases.

And finally, I believe that the FDA should capitalize urgently on the investment of the electronic medical records and implement interoperability integration with NextGen sequencing. In the United States almost every institution has now electronic medical records, we'll use Epic and will have invested a lot of energy and resources to build these programs and we should be able to access all NextGen sequencing patient's data as well as treatments and learn from each patient.

And in my opinion the long-term plan should be that the FDA should direct setting up the informational infrastructure to be prepared for the artificial intelligence revolution which will be to use NGS data, EMR data to perform algorithm analysis using artificial intelligence in decision making for optimal drug selection for more effective drugs, for more patients, faster. Thank you for your attention.

DR. PATHAK: Thank you Dr. Tsimberidou. Now if all the speakers and all the panelists could come up. On the panel in addition to the speakers we have Dr. Neal Lindeman. He is an Associate Professor of Pathology and Director of the Molecular Diagnostic Lab at Brigham and Woman's Hospital, also affiliated with Dana-Farber and they do work for the Boston Children's Hospital and as I said he is representing AMP.

And we also have Dr. John Pfeifer, he's a Professor of Pathology, Vice Chair for Clinical Affairs and the Interim Division Chief for

Washington University, St. Louis Pathology and he is representing CAP.

So first of all I'd like to thank all the speakers for, your diversity of perspectives and experience that they've shared in their presentations.

And they basically, you know, we haven't heard from Dr. Pfeifer yet or Dr. Lindeman so you know one of the things we would like to learn from this symposium is a level of evidence that you use in clinical decision making at your institution.

And we've heard that Dr. Kulkarni basically aligns with the AMP guidelines and so does Dr. McLeod's group but they have their own grading system and Dr. Tsimberidou is also aligned with that. Dr. Lindeman, can you comment on how you use the levels of evidence in terms of variant reporting and directing patient care?

DR. LINDEMAN: Sure, thank you very much. So we're in the process of transitioning from the system that we used before to the guidelines that I actually helped write along with Lia and Shashi but we haven't quite made that transition yet.

We have a five-tiered system -- it's very similar to the systems that you have heard earlier today where the 5th tier is actually the known benign variants, we don't even report those.

So we really only report the first 4, the VUS's are in tier 4. The clearly actionable alterations, whether they are for diagnosis, prognosis or therapy selection are in tier 1 and most of our decision points hinder between tier 2 and tier 3 and for us the tier 3 alterations are investigational targets, they are prognostic alterations where there isn't a very clear definitive treatment algorithm based on the prognostic significance such as the decision to transplant that was addressed earlier or they're highly associated with a diagnosis but don't establish it in and of itself.

And those have fallen to what we consider the actionable alterations, any of that -- whether it's tier 1 or tier 2, and then the tier 3's are the things for which there's investigational data, animal models, cell line studies, a pathway that can be targeted but the specific alteration hasn't been well studied.

DR. PATHAK: Dr. Pfeifer?

DR. PFEIFER: Yeah, we follow the guidelines - the AMP guidelines. One of the authors on that paper was a member of our group, Dr. Eric Duncavage so we follow those guidelines almost to the letter.

I think that Neal -- I'm going to just say that Neal pretty much summarized exactly the way we approach things at Wash U so if you want to know the details of what we do just read what Neal said and we do the same thing.

DR. PATHAK: Okay, so um, you know it's good that the AMP ASCO CAP guidelines came out with a basic framework, but basically at the point of care or in practice there's also the issue of you know, investigational targets and the use of pre-clinical data like in vitro or in vivo and you know, when you're dealing with patient care you want to ensure efficacy and also safety as Dr.

McLeod mentioned.

So I'd like to hear from everyone as to how this type of in vitro, maybe in vivo, data should be used optimally and how can you aggregate this type of information across the country as Dr. Tsimberidou just sort of mentioned so we can move the field forward.

DR. PFEIFER: So I'm going to make a comment to sort of establish a foundation for that. Because when we come to a conversation like that which is very important and I don't mean to pre-empt your question or anything like that.

I just want to establish a foundation for this conversation which I think is often ignored in this which has already been touched on a couple of times this morning and that is that NGS testing -- that NGS is a method, it's not a test.

We all, in our laboratories, use NGS methods which are these massively parallel sequencing methods that provide digital information as methods to support clinical tests that we have to do.

And I want to use an example that's come up a couple of times this morning to help underscore this conversation that we're about to have about the use of how we use it in support of clinical trials.

And that means that all of us who are at academic institutions and even those in the commercial sector, we design tests that meet a patient care need and at an academic institution those patient care needs often are different between different institutions because they have different areas of focus or different groups that are doing a specific thing.

So for example at our institution it's well-known that we're the -- now I think the second part just transplant group at Wash U. And so we have recently developed an assay which we refer to as Myeloseq which is designed around a UMI based methodology -- now just bear with me here, there's a point to this.

The UMI methodology, unique molecular identifiers, also sometimes globally known as molecular barcodes -- assays designed around that approach have a much lower false negative rate -- that is to say they're more sensitive because you use reed families and you can and so the data analysis has a much higher sensitivity.

Now it's interesting to note that this morning Dr. Berger mentioned this in the context of their battleship assay that they use across all their patients is this other assay design that they're interested in using and Shashi has mentioned it -- Dr. Kulkarni has mentioned it this morning that that's actually the assay design that they use for their tests.

At Wash U we have a test -- our comprehensive cancer test that doesn't use UMI methodology but we have a Myeloseq assay that does. Now the point is that it not only has a much lower false positive rate -- false negative rate, it has a much lower false negative rate -- false positive rate, meaning it has higher sensitivity and specificity both.

Now the point in all of this is to say that we would issue a report based on that Myeloseq assay that would disagree from what our own GPS laboratory would do and a number of other laboratories that were performing a comprehensive cancer test.

And so you would be stuck in a situation where you are saying laboratories have results that differ. Is this because there is a difference in their bioinformatics pipelines -- they're annotating variants' different, they can't find the variants in the same way.

And this raises two points -- first of all that because NGS is a methodology and not a test, different tests are going to produce different data that depending on the bioinformatics pipeline, you're going to get different answers and that doesn't mean that one pipeline for doing the annotation is better or worse than another, it just means that those pipelines are tuned to the specific tests.

And so in all of this conversation I think it's important to remember that at the fundamental level we're talking about individual tests that we're running and so as FDA considers these issues and we as a community, we need to recognize that what's appropriate for one test isn't necessarily the most appropriate bioinformatics pipeline or interpretation for another.

And the second real problem comes at the clinical side. If one of our HEME OC people got a result form the MSK tester foundation, went and got that result confused with the Myeloseq assay, they could make a completely inappropriate clinical decision.

And so that's the point -- it's sort of a two-part point and it just sets the stage that when we talk about how these assays support perhaps the clinical research that's going on at our institution, we need to understand that the assays themselves are fundamentally different in the way that they're designed or what their intended use is -- so sorry for that.

DR. LINDEMAN: I'd just like to echo what John said very eloquently right, so the design of an assay for a specific population in its medical center such as ours, at Dana-Farber is tailored to the needs of the patients and their and our clinical constituency at the Dana-Farber and it is a medical physician activity.

And two different tests can have two different results not because one of them is wrong but because of the way they're designed. And just like the same patient could go to two different oncologists and get two different treatment recommendations and that doesn't mean that both of them -- one of those oncologist's is wrong.

It's a stylistic preference, it's a difference in understanding the literature and how to apply it to each individual patient and we make that at scale at our cancer centers, but it is essentially the same kind of distinction.

DR. PATHAHK: Dr. Tsimberidou?

DR. TSIMBERIDOU: From a clinical perspective we have a large number of clinical trials so with specific targeting specific molecular alterations in oncology trials.

I think the key issue is first of all what level of evidence does someone use to design a clinical trial? How do we prioritize this trial?

For instance there are genetic alterations that carry more value or weight in determining or causing carcionogenesis compared to either alterations.

And we select treatment based on what trials are available. On the other hand I believe that we should separate where the level is very high to recommend a treatment that is more likely to benefit the patient.

And when we exhaust treatments that are likely to benefit then we can go to the lower -- use lower level of evidence for instance phase 1 trials with novel agents that have demonstrated evidence in in vitro or pre-clinical data of activity and we can -- if this is for instance, a patient with a last -- who have exhausted standard treatments or other higher or other drugs that they are more likely to benefit them than which were within all of them on these trials.

The other key issue I believe, is the lack of inadequate trials with combinations or basket trials because what the trend I see in recent years is that we have one drug against one molecular alteration for instance, and we spend a lot of time and energy enrolling patients for these trials.

Whereas, if we were able to combine targeted therapies with some cytotoxics even, or immunooncology drugs, then the probability to see a response would be more, in my opinion, higher.

So these are the novel designs we should investigate as well as basket trials.

DR. PATHAK: Thank you, Dr. McLeod?

DR. MCLEOD: I think -- I think one of the aspects that we're not very good at, at least at my institution, is providing good guidance for test ordering and the limitations and positives of a given test.

We've all seen the cases come in from the outside where their oncologist ordered an EGFR hotspot test and then wondered why they didn't find that variant that wasn't on the panel and it happens over and over again and we need to be aware of that.

We provide much better guidance around drugs partly because the margin on drugs is higher. We float our institutions on drug margin and therefore we better pay attention to it.

The testing part we kind of hope that some poor sap in the send out lab does a good job or some technician in pathology has the backbone to stand up to some high-volume oncologist -- it's just not a very good approach.

I think there's opportunity both in terms of level of evidence but also in terms of the discussion that we've just been having -- the nuances of the tests to try to make this is much better approach and that's me answering the question that I couldn't remember what you asked.

So I answered the question that I wanted to answer.

DR. PATHAK: Okay Dr. Kulkarni?

DR. KULKARNI: I remember the question so

I'll stick to that. I think his question was how do we use the in vitro level data and functional data in the levels of evidence?

So you know the AMP ASCO CAP guidelines have considered that as one of the levels of evidence so it's factored in. Of course it's not level 1 and level 1 if it's stand-alone but if it's in association with solid clinical data or a cohort of patients then it's used as an agent rather than stand-alone.

So if it's just a stand-alone then the tiers -- I don't remember exactly if it's tier 3, B or 4 or something -- it's not benign but it's just before the benign so.

DR. PATHAK: Okay that's very useful feedback from all of you and thank you Shashi and -- Dr. Kulkarni, I'm sorry and Anderson for your input and you know getting to the platform question -- that raises you know a host of very interesting follow on questions because when you use these unique molecular identifiers you can probably get down to very low allele frequencies.

And we had a discussion in the previous session about allele frequencies and how meaningful they are and whether it's possible to be over-interpreting a low allele frequency in a sub-clonal population.

I mean are there efforts to sort of codify allele frequency a little bit better or allow another sort of twist to the question is allele frequency is often a function of tumor content so should tumor content and allele frequency be carefully captured in future databases?

DR. PFEIFER: I fear that I have somehow created some confusion in here. What I meant to say by my previous comments is that NGS methods are used for tests.

DR. PATHAK: Right.

DR. PFEIFER: And questions about the significance of allele fractures depend on the tests that you have designed.

DR. PATHAK: Right, right.

DR. PFEIFER: So if you're looking at a solid tumor test in a patient with non-small cell lung cancer or liver cancer or colon cancer -- there, it's very well established that there are, you know, the genetic instability that there are copy number variations and that can change what -- the significance of a calculated allele fracture and certainly the tumor cellularity.

So in our shop we are very reluctant to put -- we're very careful not to over-emphasize the variant allele fraction. We rarely at the time of interpretation -- although sometimes you do, go back and look at the calculated tumor cellularity at the time that the specimen, you know, was sessioned into our laboratory.

In contrast, when you're using an assay that uses UMI's and you're specifically looking for minimum residual disease or very small clones, then you're right -- your allele fraction is extremely low but then you're using an assay design that is specifically looking for very rare clones.

And if you're down there around not 1% but .1% or even maybe a log lower than that. Now all of a sudden if very rare clones happen to have a little copy number amplification in there, that's not really the same -- you're not really looking using an assay to ask the same question as you are at the time that the patient presents -- at the primary presentation for a tumor or a gross metastasis, so sorry.

DR. LINDEMAN: So basically I would agree with what John -- we're just agreeing with each other here but it depends on the assay design. So some of them -- allele fraction is important and it's rigorous and some less so.

For our particular assay it is an important consideration as is the tumor content, as is the copy number assessment. We actually reconcile those in our report. So we do put the allele fraction and then we explain it in the context of the observed tumor content.

We re-review the slides on every case we do as well as the copy number. And we use that to help guide the person reading the report as to how to interpret each of the variants and whether it is reflective of a sub-population or not and we don't have the benefit of paired germlines so it also helps us in that context as well to identify likely germline variants.

And I believe that's all part of the guideline document.

DR. TSIMBERIDOU: I agree that all of these components are important and they should be listed in the report. Earlier it was discussed that they should not be mandatory. In my opinion as a clinical investigator I believe that we call all use this data to interpret clinical outcomes, particularly because patient's have multiple molecular alterations.

They have immune abnormalities or immune markers and we need to know what will be the interaction of these markers and how we will prioritize treatments.

So this is, in my opinion, a key issue in the description of the findings.

DR. MCLEOD: We certainly use it but we use it in a semi-quantitative way. So because we do primarily tumor only sequencing it does help tip the scale if one considers a germline variant and at least to send them for germline testing.

With all the caveats of (inaudible) that are present. If someone has a variant of that 37% and another one at 43%, we don’t think those are different and call it poly-clonal and -- but it can, certainly with longitudinal testing either tissue or more commonly liquid biopsy, one can look at trends in terms of a rising clone that seems to be a resisting clone and when to intervene based on that.

The one caveat with the liquid biopsies is that rarely do we get the denominator and um, that part I think, is especially important because so much of the denominator is normal DNA and we’ve even, with one well-known company, start asking just for basically reads -- how many reads, rather than the percentage because we’re wondering if someone sneezed prior to their test being taken and therefore the denominator is higher.

So you know there’s some work to be done there but it does have some utility.

DR. PATHAK: Dr. Kulkarni?

DR. KULKARNI: I'll be very brief. I have not a whole lot to add but I think with using UMI's we have extreme high sensitivity than what we know what we can do with it.

And again, understanding the complexity of the biology itself and technical limitations, we do put the allele frequency in the report to be used later.

And of course we use this internally to have -- understand the integrity of that particular run looking at copy number changes and all that. But I would add that as we are starting to curate clinical grate variance in ClinVar using the ClinGen somatic workgroup platform, we are asking the investigators to put weight on allele frequency and the type of -- a little bit of detail of the type of assay used, you might not.

So I think as we build this repository with outcome and more clinical data, we will -- we will know more how to use this.

DR. PATHAK: Yes, thank you so much. So um, that was very sort of insightful, illuminating input from the panel on that particular question which came up earlier today.

But you know one of the things that I am curious about really is how does the actual variant interpretation process occur after institution?

Do you use informatics pipelines that you developed yourself, do you use databases -- what is the role of molecular tumor boards and how do you sort of integrate this information when different databases may disagree?

DR. LINEMAN: So maybe I'll take a stab at that first. So for us we use the databases as a resource, not as part of a device. And so I like to think of the database kind of like a textbook, just a very highly searchable one.

And it's a repository of information and there are several of them and just like there are several textbooks that you can use and the interpretation that we do is done by our physician staff so we're a little bit fortunate I guess.

We have 31 physicians on our faculty and we have several on every day and that's senior faculty and then we have our trainees, our residents and fellows and we have some doctoral scientists.

And we spend a lot of time going through each report similar to Mike's about 150 or so a week and we incorporate the signal data that we get off the analyzer with the medical knowledge with what's found in databases and literature searches as well as since we're at one institution, the medical record and accessing what's going on with each patient.

And we make a customized report for everyone. We store that knowledge in a knowledge base which we have been building up over the last 5 years so we can pull back in our previous interpretations but then we modify them as necessary for each individual case.

DR. PFEIFER: So we do very much the same thing -- geez Neal it could be like one of us and not both of us here. But we do the similar -- a similar process.

There are some minor differences in that we have a bioinformatics pipeline that essentially pre-templates the report so we have this database that we draw from that draws from all the publicly available databases and we have some licenses for others and then all that information is pre-templated into the report based on the various tiers of evidence.

Every single report though at that point then -- as Neal's phrase, goes in front of a faculty member who has sub-specialty boards in molecular pathology -- whether it's through the ABMGG pathway or the American Board of Pathology pathway.

And so all of those variant calls are reviewed -- not that the variant calls interpretations are reviewed to make sure that they're appropriate in the context of the patient.

Now this is a very labor intensive -- very labor intensive process and one of the reasons we do it is because we are all after all an academic institution and we have trainees and this is part of our process.

It is interesting to keep in mind that we have trainees that are at our institution and so every case when it comes to me has already been reviewed by a trainee or what we call these variant scientists -- people who are masters or PhD level who pre-template those.

Trainees have this um, wonderful habit of changing the level of a variant call in a way that I wouldn't have changed it but causes me -- forces me to think about things that I wouldn't have thought about on my own.

And they create a lot of inefficiencies but they also ask a lot of really interesting questions. And so at the end of the day, if it's not entirely clear what should happen, we sit down as a small group and discuss why do you think that should be a level 1 versus a level 3 or a level 4?

And we -- as they do in Neal's shop, after we come to a decision we write a comment and that is stored by our informatics pipeline using natural language functionality so the next time we come up with a patient who has thymic carcinoma of squamous morphology, et cetera, et cetera, and a mutual kit, that is automatically offered up in the pre-templated version of the report as something we could draw from, so we're aware of that.

DR. PATHAK: Yes, so another part of the original question was -- that I asked, which actually which I forgot to mention, was how do you use your internal databases and your internal experience you know, as you have just mentioned?

I mean even for the quality of NGS calls?

DR. MCLEOD: Our health research informatics database has just over 600,000 patient's worth of data in it in which a small subset has had molecular.

But it's been very useful in that we can -- first of all we can look at the tens of thousands, not hundreds, that have had the molecular and ask questions about have we seen this before at this institution?

Have we seen it at some other institutions, et cetera? If we've seen it at our institution we can then go straight to that data and understand, okay in the 14 people where this happened, here's what they received, here's what happened.

It's anecdotal and 14 people is barely a case series but it at least gives us something to look at. We also have more of a therapeutics orientation at that stage so typically what we're looking at is what we call VAKS -- a variant of almost known significance.

So it is something in a domain that matters but we have never seen it before. And so then we even get to the point of looking at in rare cases, the protein docking information -- all this kind of stuff that we don't believe should drive therapeutic decisions but could allow us to have a hint at whether we should even think about this therapeutic.

But that's you know, again, down to the tie-breaker where we're happy to have anything. You know something is better than nothing. Often is where we get to and the databases will help with that.

DR. PATHAK: Thank you.

DR. TSIMERIDOU: I think the tumor boards are very important to decide -- determine which treatment a patient should receive. And they should in our institution, of course we have the precision oncology decision support team that provides the reports based on -- with molecular profiles done, NextGen sequence system was used and then we have interpretation of all things and it matters to us as I shared earlier with clinical trials if they're available.

But I think what is important is the move from this objective selection of treatment that happens in clinic by one oncologist to the more objective that will be reviewed in a multi-disciplinary fashion by several experts in the field and in my opinion is very important to incorporate the purview of the expert pathologist, what was missing, in my opinion's, expertise.

So we have to have experts in the decision-making process in precision medicine. The other challenge I have seen recently is that with um, FDA approvals -- for instance the approval of fludarabine for MSI hypoplasia.

We have to be -- to incorporate these changes very quickly in our practice in order for instance, MSI status for all patients because automatically these patients have approved fludarabine and we should keep up and make sure that our systems are updated continuously with what is FDA approved because we need -- when we're discussing using molecular profiles and different levels of evidence it would be, I think, important for a patient to receive therapy approved treatment and then select the investigation treatment based on profiles.

DR. PFEIFER: It's an interesting question as to where the line ends between the technical part of the testing and the practice of medicine, if you will, begin.

One of the things that I was fascinated by in these excellent talks this morning -- I keep looking at Mike because he happens to be in the front row here is -- I think we're all -- the ideal here is that all of us are drawing from the same or similar databases.

So for these variant -- you know, these level 1 or level 2's that we all agree on those and that our software comes up -- can find those -- our bioinformatics tools and that they're annotated in the same way.

And by annotation I mean that the variant is annotated according to standard um, standard procedures so that we're all calling the same variant the same thing so we could find it and we know that this is not necessarily a straight-forward thing sometimes.

Then it's -- then the annotation includes what drugs that this is responsive to in what tumor types and so we're getting that same level 1, level 2 and maybe level 3 right -- that's fundamental.

I personally, sometimes I'm a rather cynical guy and I wonder if that's actually happening out there. I mean we're all drawing -- we have our own bioinformatics tools -- that is the first goal is to make sure that that's happening.

That, to me, is a component of the test. You want the test to -- that test report at the initial level -- we call it the templated report, to have that right -- to have that right.

Then there's this question about once the report is issued, where do different institutions draw the line as to what is also included in that report?

At academic institutions that you've heard us talk about that includes some component of the practice of medicine -- interpreting what that means in that patient in conversation with our clinical colleagues.

Well that perhaps muddies the water as to where the test ends and the interpretation of the practice of medicine begins. In other laboratories it might just stop with what we would call the pre-templated report, although even if it stops there that's fantastic, but we should all get to that point correct.

We should all be exactly the same at that so I don't know whether that's helpful but.

DR. PATHAK: No, no, absolutely. So you know, it's great that there's, you know, emerging institutional knowledge at every institution or most institutions.

But one of the problems I see is making sense of rare variants. And I mean, I think that Dr. Tsimberidou mentioned earlier that there should be some type of effort in terms of gathering together all this information and sort of, you know, maybe like a meta-database or something.

And have you been -- this question is for everyone -- have you been actively participating in these type of activities, like contributing data to databases and registries and what's the utility and how can we sort of harmonize the system so that we could produce sort of the harmonic output that Dr. Pfeifer sort of alluded to?

DR. KULKARNI: I think I will take a stab at this. So yes, so we are firstly our institution is contributing to ClinVar database and I think we are getting more and more partners from industry and academy to participate in the ClinGen somatic workgroup.

I think that will drive this much further. Rather than just focusing on the standardization of the way the data are interpreted -- at the end when it's in the database I think we have to do a lot of work upfront.

So the way we do the workflow in the interpretation is similar to what my friends and colleagues have mentioned, you know -- faculty members, board certified pathologists, geneticists, fellows and all of that.

In addition to this -- because we had a large clinical lab with 150,000 test volume per year, we just could not take this and scale it up because it's such a labor intensive process.

And we do extreme complex testing like whole exome sequencing and germline and all that so in order to make it scalable we have created a new division called "Clinical Genomics Interpretation".

We have about 20 PhD level scientists -- we call them clinical genomic scientists and we have standardized SOP's and so we are trying to make it very standardized and we don't want to have two separate -- just like John said, you know, where does it stop to being a technical interpretation versus the practice of medicine?

I think in our organization practice of medicine starts when the board certified person is signing out any changes. But before that I want to make it standardized and run of the mill and very scalable. So we are putting those SOP's together. We have done quite a lot of work in not only germline but somatic also.

In addition to doing this we are partnership with Mayo College of Medicine and Mayo Clinical Labs to come up with standardization. In about a year or so we will have a conference in Houston where we are all going to get together.

We'll have workshops where we will be giving in silico data to all these people. Essentially this is a new career path for a lot of people -- clinical genomic scientists, ovarian scientists, but there is no standardization or board certificate for that.

So we're trying to imitate some of that through partnering with big labs who do similar kind of work so I think it will be a start but, you know, just you know, needs more work.

DR. PATHAK: Okay, unless someone wants to speak I think we need to take questions now. Does someone say any thoughts or final words on this?

DR. TSIMBERIDOU: I would just like -- I would like to add that at the end of the day we're going to have access to drugs because we do all of this interpretation and very detailed reports.

And then you have a list of 20 genome abnormalities and you have -- you will be lucky if you have one clinical trial or one drug that the patient can have access to and that their insurance will approve or the trial will allow, based on reasonabile criteria.

So there is a lot of work also to do from there forward to get the drug to the patient.

DR. PATHAK: Sure.

DR. LINDEMAN: I would actually like to say one thing. So the question was about rare variants and I just want to make the point if it isn't obvious and someone joked about who my daughter is going to date.

But there is a difference between the germline space and the somatic space and the rare variant problem -- and it is a problem, is rare in cancer for a certain reason.

And so it's just important to understand that the infrastructure that applies for inherited disorders is a little bit different from what's really relevant in cancer.

And most of the calls that these assays make are pretty clear, and quickly and easily bend into these most actionable categories.

DR. PATHAK: Okay, thank you so much and we have time for questions now if anyone wants to ask questions?

DR. MCLEOD: Maybe questions about the lunch menu?

DR. PATHAK: Yes, please introduce yourself.

ANICO: Hi, my name is Anico and I have a question on this manual creation which feeds into these databases even though you know, there are different levels of checks that you can implement.

Should we worry about, you know, people make -- human beings making errors creating these databases? And I think one of the um -- follow-up on that would be I appreciate that there are some committees that review the level of evidence.

I'm assuming that also effects the turnaround time so I was wondering if you could comment on those two?

DR. LINDEMAN: Yeah, turnaround time -- so that was our big initiative last year was to try and turn a six week test into a two week test which I'm glad to say we were able to do, but it wasn't easy.

And it came basically by putting an awful lot more people on the project as well as some other process redesigns. Can a human make a mistake? Sure. A human can always make a mistake.

We have a little bit of a gauntlet that each of those curations go through so it's got to get through actually three -- I didn't describe it completely, but three sets of reviewers before it goes out.

We have embedded our faculty. We have a faculty member with each of the disease centers at the Dana-Farber so we don't rely exclusively on the tumor boards which Howard kind of made me chuckle because they do tend to be really kind of look at this weird thing next.

But those embedded relationships are critical so every oncologist at our center has a member of our team that they have a personal interaction with and connection with and we go to their meetings every week.

And then as we encounter those same variants we do call them back up and sort of pre-populate and we review them. And if we see there's a mistake, there's an opportunity for the next person to edit it and go back and contact if a mistake was made or more accurately really it's sort of medical knowledge that's changed.

And so the meaning of the study and the variant may have altered over time, but that's the process we have.

DR. PATHAK: Dr. McLeod?

DR. MCLEOD: I think part of it also is many of the institutions here or maybe all have training programs associated with them and so we have the benefit -- we have a personalized cancer medicine fellowship that's mainly medical oncologists, pathologists of various types and clinical pharmacologists, all of the same fellowship.

And so we have different lenses being viewed at these and they go back and as part of their early training part portion review past decisions. And it's -- would we still make that decision today and if so, what would we change et cetera.

Also, as John alluded to we also have a point where if we've seen it before -- the language pops up we can then say alright, is that still true?

And so that's a component of it but humans are involved. Although I would remark that when computers were involved with the different AI approaches that have been taken to date, IBM Watson being the highest profile, they have not done better and have done worse, in my personal opinion.

So we're not there yet. There will be a day when AI, when HAL takes over and we're all working for them but we're not there yet.

DR. LINDEMAN: Actually in my experience the biggest mistakes are the informatics errors that sometimes happen behind the scenes in a systematic way and it's the manual process that catches them.

When the human makes a mistake, it's usually one at a time but when a computer makes a mistake it's much larger scale and so I'll just add that caveat for anyone running a lab -- know what your informatics is doing.

DR. PFEIFER: Yeah, we have, at Wash U in our shop, we have worked very hard and Neal mentioned this -- there is this gray zone or there tends to be this gray zone between where the technical part of the test ends and the practice of medicine part begins.

And the interpretative piece to me is where the practice of medicine begins. We've worked very hard to make that core bioinformatics part of the technical component of the test.

We do not have time to go through in every single case and make sure the knowledge base is right and make sure that those pre-templated variants are correct.

We like everybody who runs a major lab, has invested a lot of resources to make sure that that point is right. I don't want anybody to be left with the impression that we're going back every single time we see a G12D and KRAS that somebody's going through the literature -- no.

All that stuff is pretty templated. We take -- we work very hard to get it as far down the path as we can so that the work that we do as far as that interpretative piece, what does that mean in this patient?

Why would we move this variant from a level 4 to a level 3 or maybe a level 2 so that we can actually dig deeper and make those finer calls, but I don't want anybody left with the impression that we're manually doing a lot of that work.

These pre-templated reports, the things that come to us are very sophisticated and very advanced so that we can spend our time on the really detailed clinical questions that are clinical colleagues have called on us.

You know when you see Mrs. Smith's stuff I want to know about this or that, so.

DR. PATHAK: Well thank you. We're running over but we can take the last two questions.

MR. ABAAN: So it's obvious that you all have really intricate workflows for running the evidence and making the final decisions. But I want to ask something about sharing the evidence between groups -- you know, besides ClinVar, when you find a novel discovery -- let's say there's a new variant that you found, you know, you dig in you found something you acted upon, you collected the outcome and said, yeah this is good, it works.

How does the sharing to the community work? Do you have a mechanism to say oh, you know, this worked out really well, you know, besides just publishing a paper and putting it again out there that somebody has to go dig and you know, bring it back into their knowledge base?

Because you each have your own knowledge bases but how does that knowledge get shared across?

DR. PFEIFER: So we used to do a really good job sharing and then HIPAA came along and our compliance office has really, really, really, made that difficult for us to do currently.

DR. MCLEOD: And, there are components of our health systems that think everything is worth a billion dollars and so if we give it away we're giving away the shop and that's a problem.

Most of this is paid for by the clinical enterprise not by the NAH and so there's a different level proprietary view that can come no that and it is a real friction between data sharing and method sharing in that context.

DR. TSIMBERIDOU: At MD Anderson we use it internally to make decisions about the level of evidence, the data we generate. But as we discussed, because of contracts with pharmaceutical companies and other you know, entities, we are very limited so we cannot publicly state otherwise other than you know, publishing articles and presenting them at the conferences.

MR. AUDIA: Niraj Audia -- it seems, looking -- keeping in mind at the end we need a companion diagnostic or the end result will be a companion diagnostic.

With so many touchpoints, how do you envision the final diagnostic to look like? Is this limiting us down to having a centralized model that you just have a CDX offered FMI or Sloan Kettering or wherever, or how do you enable this in a commercial setting to lead to a more democratization of availability?

DR. LINDEMAN: So my personal preference would be rather than a companion diagnostic, have a companion analyte and specify the performance characteristics that a test needs to have in order to detect that analyte properly and not restrict it to a specific methodology and a specific assay.

So to say you need to detect the BRAF 600E at a certain level of sensitivity and accuracy and precision and that should be the determining factor.

DR. TSIMBERIDOU: I think the single companion diagnostic for one drug is not a functional or effective process. In my opinion, as an oncologist I would like to have panels across tumor types and screening as many alterations as possible because at the end of the day we cannot predict how many -- what alterations a patient has.

In 2007 when I started the personalized medicine program at MD Anderson we were able to order only one or two alterations. For instance, BRAF for patients with melanoma or KRAS for lung cancer -- and we had an issue with a tumor so we could not -- we didn't have adequate tumor to order all the alterations.

Now, with so many drugs available and as we move forward in incorporating multiple data and not only for targeted therapies but also immunotherapies and other programs that are in development -- in my opinion, we cannot focus or stay on that model.

That model does not work for patients or healthcare providers who have to move to large panel approval across tumor types and to be done as early as possible so we can learn about patient's tumor biology and offer them the best treatments possible.

DR. PFEIFER: I totally agree with Neal on this. My problem with the companion diagnostic model as was just very nicely articulated is that process takes too long.

By the time that that approval comes along the field has moved on. The other problem with the companion diagnostic is it starts in my mind, mixing up NGS as a method with NGS as a test.

As Neal said, what we need is to recognize that we're using these massively parallel sequencing methods to run these clinical tests and the clinical scenario in which these tests are being ordered changes rapidly and so we're doing this Myeloseq assay which there's real concern internally is going to put our comprehensive cancer test out of business.

And I keep saying to people, they are providing completely separate pieces of information -- completely different levels of sensitivity and specificity.

So when an oncologist is confused, you know, about which of those tests to order, they think they're in competition -- they don't understand what we're doing. So my problem with the companion diagnostic is it has to be very clear that that has a very limited clinical, you know, utility and that the people who are out there -- our clinical colleagues who are ordering these tests, need to be aware that an NGS test is not the answer, it’s the test that’s designed to your patient’s need.

And that’s like the second or third time I’ve said that which to me is an indication that I really don’t have anything to say to contribute to this panel.

DR. PATHAK: Well, um, I’d like to thank the panelists for the very interesting, illuminating discussion from multiple perspectives and I think we can break for lunch now. Thank you.

(Lunch break)

DR. MADISON: I want to thank our morning speakers for a wonderful first two sessions, our moderators for providing some really good sets of Q and A for everyone that participated in the question and answer session from the audience.

We appreciate you taking some time to provide your input as well. As a reminder my name is Hisani Madison, I’m a Senior Scientific Reviewer in the Division of Molecular Genetics and

Pathology in CDRH -- Center for Devices and Radiological Health -- and I'll be the moderator for session three which is: Best practices for use of public and private databases for variant classification and interpretation in oncology.

And so our first speaker for today is Dr. Heidi Rehm. She is Director of the Partner's Laboratory for Molecular Medicine and an Associate Professor of Pathology at Brigham and Women's Hospital and Harvard Medical School.

She is a leader in defining standards for the interpretation of sequence variants and a principal investigator of ClinGen, providing free and publicly accessible resources to support the interpretation of genes and variants. Join me in welcoming Dr. Rehm.

DR. REHM: Thank you, it's a pleasure to be here. This has been a great conference so far. So I just have two disclosures -- I receive NIH funding for the ClinGen program and I am employed by labs that offer fee for service genetic testing at Partners and Broad Institute.

I also just want to point out my general focus is more on germline diseases and I'll show examples that span particularly germline examples, but all of the principles I will talk about are applicable to both somatic and germline cancer which is obviously what we are talking about here today.

So I just wanted to start and think about the different data sources that we use to classify variants and where that information comes from. So I sort of divided things by publications, databases as well as clinical data that we get from healthcare providers.

So to just think about some of the different benefits and drawbacks of each of these sources -- so it's actually hard to point behind me. I'm over here.

So for publications they're peer reviewed whereas databases don't have a lot of peer review. On the other hand, a lot of publications sit behind access fees and firewalls.

So it's difficult sometimes for the public to get access to papers and the research is often transient and sometimes you can't even contact whoever is listed as the contact person on that paper.

The peer review process is also highly variable -- almost never do the peer reviewers actually dive into the evidence on individual variants and whether they were classified appropriately.

The data is largely unstructured with limited clinical info on each of the variants -- it's more aggregate information. The variants are generally not QC'd for nomenclature and also it's difficult sometimes to discern a case that's present in multiple publications but is in fact the same case.

Yet the data is indexed in PubMed and is generally searchable except sometimes when those variants are deep in supplements.

For databases -- has the benefit that a lot of the data tends to be more structured and standardized in its fields -- the variants are typically QC'd for nomenclature; there's a much lower barrier to submission -- your work doesn't have to be of broad interest and impact to get into a publication; and you could also share small bits of data one piece at a time which is more difficult in the publication arena where people may wait years to amass enough data to publish it.

Drawbacks on the databases -- there is also sometimes a fee for access, like the HGMD database and the inclusion of supporting evidence or an interpretation may vary in those databases.

Clinical data obviously, is incredibly useful for classifying variants but the quality of that data does depend on who's providing the data and in what format.

I've had rec forms filled out by administrative assistants and other things like that that are just incorrect. The other thing I think, in general, and this has been mentioned by Howard and Neal Lindeman and others -- a lot of the variants that we deal with are extremely rare and does require a global sharing to amass even small amounts of data to classify variants.

Particularly, in germline cancer -- as Neal mentioned, a little less so in somatic -- but we still see lots of very rare variants even in somatic disease.

And so there are very few variants that really have enough data to take statistically robust approaches to interpretation to be able to do really large case control studies with very well validated functional assays.

A lot of what we deal with are small bits of information that we have to use to classify. So I'm going to walk through a few of the different databases that are in use today to get a sense of what's out there, what some of the benefits and drawbacks of each of those databases are.

So gnomAD and ExAC are two of the very large, incredible useful public databases of allele frequencies. This data has been free since 2014 -- data coming from over 138,000 exomes and genomes.

There are very robust allele frequencies, there are sub-population frequencies, they have excluded severe pediatric disease and they also have a version that excludes cancer cases, specifically for the cancer community to use, and there are accessible quality metrics and views of the raw data that you can see on an individual variant level.

A major drawback though is that you largely cannot get access to a phenotype data on an individual case. But in those databases there is over 277 million variants in gnomAD and ExAC combined.

Human gene mutation database for germline variants has been around for a long time -- I think since 1993. It’s probably the best source to identify germline variants reported in the literature. However, it comes with a high cost to get access.

And generally the curator simply enters the claims from the literature which are often not correct. This is just a figure from one of the papers that I published where for healthy genome analysis only 8% of the variants reported in that database as pathogenic actually had evidence to support that claim -- so a huge issue with the direct dump of data from the literature.

From the 2018 stats, there are over 220,000 variants in HGMD. The Leiden open variation database has been around for a while as well -- since 2002. It has the advantage that users can actually set up their own instance of this database and put their own case level data into it and it actually does a reasonable job of allow you to track a basic data at the case level and then aggregate that up to a variant level.

The drawbacks are -- it's highly variable in terms of the content for any given gene, some are just completely devoid of any data, others limited, if somebody took on that role, are well curated -- so it's very variable. It's also difficult to get stats on what's in there because they added in a huge amount of genome and exome data that is sort of polluting the data on variant interpretation.

ClinVar has been around since 2013. This data comes from many sources -- clinical labs, researchers, databases, other databases, clinics, patient registries -- the majority of it is from clinical labs, about 80%, and that data is reasonably kept up to date although the research and literature data is less kept up to date.

The system does distinguish by review status using that star system which I'll talk about a little bit later. The drawback is there are fewer structured submissions of case level data -- it's more of a variant level database with summarized evidence.

And the quality of the interpretations does vary, depending on the submitter. And supporting evidence is not present in about 19% of the entries.

As of this weekend there are over 375,000 unique variants from 180 submitters from 63 countries -- so very widespread use at this point.

I'm less familiar with the somatic cancer databases but there's a project called VICC, the

Variant Interpretation Cancer Consortium that has been working to bring together a lot of these different databases.

You've heard about some of them this morning from different groups that are all working in the cancer space. One of the challenges with these databases is that they use a variety of custom nomenclature and ontologies. This is an example of the same variant representing three different ways and three different databases.

And so this group has been working to sort of bring these structures together. This effort began just in 2016 and they've now gotten 8 of the knowledge bases committed to share data and integrate it, 6 of them have been integrated so far and are accessible on this website, and representing over 17,000 variants that have been curated for somatic cancer.

And that is all freely accessible. You can get both downloads, API access, and they're also submitting entries to ClinVar and they've normalized the data structure now to the AMP guidelines, so that's a large effort bringing some of the somatic cancer databases together.

If we look at what's in these databases now and think about how to use the information in there -- so this is just a figure from ClinVar looking at the top 10 genes in terms of numbers of variants.

And the three different bars represent the total unique variants, the variants by classification, and then conflicts. So using our ClinVar Miner tool we can look at conflicts in those data and the three different tiers, confidence, so pathogenic versus likely pathogenic which in ClinVar is not recorded as a conflict.

The orange is not clinically significant but VUS versus likely benign, and then this small little red tab there is the clinically significant variant.

Some of this is pathogenic, someone else says VUS likely benign or benign, so it gives you a sense of how many conflicts relative to the percentage of data is in there and those are an area of work for us to resolve.

The other question is where do those conflicts come from? This is a nice paper from the Invitae group where they looked at the source, the type of collection method -- was it clinical testing literature from publications, research or curation efforts?

And you can see most of the conflicts do come from the published literature and a smaller set from the research, much less from clinical testing labs, particularly for cancer genes. A lot of the literature is not correct and represents outlier interpretations.

So overall the concordance in ClinVar, particularly for the medically significant differences, is quite high in ClinVar, but we still have the opportunity to resolve the smaller set of variants that are discordant.

And we've been doing projects to take the data and resolve it. We've been -- this is one of the early publications where we showed 87% resolution of the differences of interpretation that were in there.

We've now scaled this project to encompass every clinical lab that's submitting to ClinVar in the germline space and we're doing an outlier approach where the lab that is the outlier, if there's two-thirds of the majority to agree, then is asked to first review the variant -- there's an auto advance going on here -- so they'll be sent the variant for review.

And that process allows us to resolve the majority without all of the labs having to reassess and then the remainder -- the 37% from this effort, then we share underlying evidence and work to resolve it after sharing data. So this is a major effort that's underway.

One of my goals and our goals as the ClinGen consortium is to get more evidence specifically in the database as opposed to having to call up a lab and ask them to send it.

80% of the entries today do have the supporting evidence in the database and you don't have to call them up and ask for it although most of the labs that I talk to do sent it to me but that is a barrier.

And so we'd like to get more of that data. Some of the challenges is just harder to submit your case level data -- other labs feel they need specific consent to do that -- so we've been developing strategies to try to get more case level data in there.

One of them is using our Genome Connectpatient registry where patients actually agree to share their clinical reports and then we submit the variants to ClinVar, they fill out health surveys and all that data is then submitted to ClinVar, so we've done a lot of that submission.

There are other clinics that are now submitting paired data, so their clinical lab submits the variant interpretation and then they submit detailed phenotypic information.

So this is from Geisinger and you can see the very detailed clinical data that's being submitted into ClinVar with the interpretation that might have come from a different clinical lab.

Here's another example from the Stanford Center for Inherited Cardiovascular Disease where they submit their interpretations to ClinVar and I've highlighted in red some of the detailed phenotypic data that they add in to that interpretation.

And this is actually a variant that's an example of, you know, there are six interpretations -- all different, ranging from VUS to pathogenic. By aggregating all of the data in the ClinVar database together, the expert panel was able then to classify that as pathogenic by bringing the detailed case level evidence that's presented in ClinVar together.

Another example of direct patient sharing clinical data -- this is a patient who was pregnant and got a genetic test that revealed a VUS.

Very concerned about her pregnancy. They contacted my lab because I was one of two submitters that submitted a VUS interpretation on that variant as did GeneDx but we were able to combine the data between GeneDx and my lab as well as the data from the family and then reclassify that variant as likely benign -- it keeps advancing by itself here.

And we were able to then resubmit this variant entry back into ClinVar with the evidence. The family actually then signed up for Genome Connect and put all the father's, you know, phenotypic data from -- who also had the variant, he wasn't healthy -- and that would all be able to be shared and so this is our current ClinVar entry with that data in it reclassified.

Here's another case -- this is just last week -- we assessed the variant as a VUS but then we noticed there are three other labs submitted to ClinVar who all agree this was a VUS -- Ambry, GeneDx and InVitae -- but then we requested detailed case level data.

We got 5 cases' worth of data, all patients with colorectal cancer and one with breast cancer. We also had two examples of abnormal immunohistochemistry data showing an absence of

PMS2 really critical, strong data in terms of the functional side.

Using the ACMG guidelines we were then able to classify that as likely pathogenic based on that detailed case level sharing -- so just really good examples of using that evidence aggregated together from public databases to classify.

Another effort to try and encourage additional evidence -- detailed evidence sharing -- so we at ClinGen have launched we call “The lab List”. This is a set of laboratories that meet a minimum set of requirements for data sharing to support quality assurance. Labs go and fill out the survey once they think they’ve met the requirements and then they can be posted on this list showing that they meet the primary requirements.

We also give them badges for additional requirements that they have met, so whether they submit the supporting evidence with their entries, whether they’ve actually submitted five years’ worth of data, whether they participate in discrepancy resolution and whether they have direct patient consent mechanisms for sharing additional data.

The supporting evidence being submitted into ClinVar will be a future requirement to be on this list and this list then allows providers, insurers and others to determine who they may order tests from or reimburse based on whether they're adhering to certain basic requirements for data sharing, so that has increased the amount of evidence and submissions going in from labs that are working to be on this list.

I was also asked to address the star level in ClinVar and how one uses that. So as a submitter your submissions can go in with either -- no stars, one star, three stars or four stars, depending on what criteria you meet.

So if you provide your criteria and your evidence or being willing to share it upon contact, you get a single star for your submissions.

If you are an expert panel that has to be approved by ClinGen, and these are a lot of the panels that have formed -- the dark green ones are currently approved -- then your data goes in as expert panel classified and then there are also the practice guidelines the CPIC guidelines are going in at that level. ACMG classified CF variants are at that level. So these classifications trump these which trump these and so on and that's sort of how the star system works.

I should note that ClinVar is thinking of getting rid of the star system -- or the stars -- because there's a lot of confusion about what stars mean. Some people think it's more pathogenic which is not what the stars mean and so they are likely to move towards descriptive terms only, probably something similar to the terms that are actually written out next to the stars like practice guideline reviewed by expert panel, et cetera.

And this is the -- when the individual submissions go in at this level then the overall variant status will be described as these terms depending on how many submitters and do they all agree and things like that -- so that's kind of how the star system works.

When you use these review levels in ClinVar -- whether they're labeled with stars or terms, it doesn't really matter -- generally we find them useful for filtering variants as well as perhaps prioritizing when to follow-up with classifications that you may disagree with.

We don't generally follow-up with the no star submitters but we do with the single star and above. I should remind everyone, however, that expert panel interpretations can get out of date.

There's a date there on every interpretation and evidence continues to amass and so as those things get out of date you have to think about whether they're incorrect.

And just because labs all agree with each other doesn't mean that it's correct, it can still be incorrect or out of date.

So all of the high quality clinical labs always review the actual primary evidence and determine if new evidence is available and take that into account.

Another take home message in general about using data sources -- really think of them as data sources not correct claims. You should use the actual evidence and take into account possible concerns about quality and get to know your data sources.

And the claims must always be assessed -- reassessed with the total body of evidence and also keeping in mind that no single data source, public or private, is ever comprehensive. Each one has different data than the next.

So I'd just like to acknowledge all of the labs, clinics, patients, researchers and organizations who shared their data in these databases, the many members of the community who create the databases to share data and curate variants to improve our knowledge and we are always looking for volunteers in our various efforts.

So if you're interested, come see us. Thank you, I think we're going to move on to the next speaker.

DR. MADISON: Thank you Dr. Rehm for that wonderful talk. Our next speaker is Dr. Shaw. She is now the Executive Director of the Khalifa Institute for Personalized Cancer Therapy at MD Anderson Cancer Center and she's here today to discuss some of the observations related to use of public databases rather, as a representative of AACR's GENIE Project, thank you.

DR. SHAW: So um, I have nothing to disclose today and today in this talk I'm not going to be discussing off-label use, however if you guys ask about that study that we did with the AACR GENIE data at the end, then I will be, but it depends on what you ask me.

I'm going to be talking today from the perspective of a user who tries to help clinicians make decisions when they get a report back, particularly from our or a commercial provider of sequencing data -- when they get it back in their hands, how they deal with that information.

And how to then utilize that information, where they get it from in the public databases that we generally use to annotate our cases, how we use those data to interpret and make decisions.

So I'd like to start with this set of slides because a couple of months ago I spoke at an FDA session similar to this but about the upfront or front end of the sequencing pipeline and how we kind of joked that theoretically this should all be relatively simple. We talk about how we kind of have this under control. We have a patient population with a set of detectable -- we presume are detectable -- biomarkers, we just have to get a sample of their tumor.

We just sequence it, we tell the doctors what's there, we interpret with some algorithms right and then the right patient gets the right drug and that's precision medicine -- patients do better and we're done.

Unfortunately of course, that's not the case. Many times we can't even detect the biomarkers either because we're using the wrong material, the sample is too old, we're using the wrong technology, because not all -- just because you use NGS does not mean you're doing the right assay as we discussed in the last panel.

I personally believe that we should be doing both tumor and blood in order to make sure we're getting the right calls for our patients and a single analyte assay from somatic only is not necessarily the optimal assay for most of our patients.

Are you doing amplicon or hybrid capture based? What are the inconsistent terminologies that we all utilize all the time and so the same alteration in two different reports might be reported differently just because of how they're termed?

What are the different algorithms we're using? And that all those different filters ends up leading to only about 10% of the patients getting the right drug, that means the matched drug, most of the time.

And so the last time I spoke on this I talked about this upfront -- this is the series of steps in getting to your clinical report in the electronic health record.

You know a lot of people talk about the problems with the sequencing and how if you have a bunch of labs that do a bunch of sequencing, not all of the mutations that they call are all the same.

I would argue exactly what I think Neal argued that I think we can deal with that through the reference materials and making sure that we deal with establishing our confidence in known variants.

But actually what I think the wild, wild, west is not so much the calling of mutations but once the mutations are called how do we actually interpret those mutations and what they mean for each individual patient at every moment during their clinical care paradigm.

Because those interpretations absolutely change depending on where that patient is when they walk in your door. Is this that diagnosis? Is that that first recurrence? Is it the 17th recurrence after they have already been treated with half the drugs that you've listed in your panel?

So what the GENIE Program did -- and so the AACR's GENIE Program is hopefully something well-known to this audience. This is currently 8 but now just expanded to about 17 different international cancer centers all over the world. At the point of this analysis we did 8 different centers and we had put 20,000 patients of CLIA certified or equivalent ISO9000 certified sequencing data into the public domain and three of these groups -- Vanderbilt Ingram Cancer Center, the Memorial Sloan Kettering Cancer Center and MD Anderson Cancer Center --we all have knowledgebases and we all put all of our knowledgebases kind of in terms of how we classify variants and what we would consider actionable together and we try to figure out of the GENIE population, what patient -- what percent of those patients would be actionable?

And we did this for the simple exercise of trying to understand what our baseline is, because when groups like MSK just published their impact paper, they suggested they matched about 10%.

MD Anderson matched about the same percentage -- 10% essentially. So that sounds like we're failing, right, if you see a number like 10% it sounds a little scary, but ultimately just at the gene level with 80,000 mutations across 20,000 patients the absolute best we could have ever gotten was 32%, so one out of 30, so we're actually doing fairly good when you consider that when we match 10%, the other patients, right, are not matching because of the fact that this is their third tumor, they're not qualifying for clinical trials, the drug is not available for compassionate use, et cetera, so we have to take that all into account.

So what we did with these databases, we looked at these different levels of evidence essentially for what we would consider actionable and compelling.

And what I want to point out is, why did we even have three different databases? And what you've just heard about is Dana-Farber also has theirs and probably every other camp -- Baylor also has theirs right?

So every one of these academic hospitals are doing this themselves and the question is why? Why aren't we just using public databases or why aren't we using commercial providers to do the service for us? It doesn't maybe seem that hard to call a BRAF V600E mutation and in fact a BRAF V600E mutation in melanoma and a couple of other tumor types is actually pretty easy and I would argue we all probably get that right.

But I'm going to go through a couple of examples that are real examples from real patients that happen to be MD Anderson cancer patients because that happens to be where I work, but argue to the point of why we are all doing basically this reinvention of the wheel so to speak and why we have to be careful when we go to these public databases and we're not necessarily sure of what rules are being applied to our cancer patients' reports.

So this happens to be an MD Anderson cancer patient report. I've obviously de-identified this report and I've taken off who necessarily actually -- it's kind of obvious who created this report.

But this is real MD Anderson cancer patient report, this patient did not have any alterations that were considered actionable on page 1, and this is esophageal adenocarcinoma. However, on page 23 of this report which of course none of my clinicians actually -- I don't know maybe Lia, maybe Lia would look at page 23 but most of the clinicians that we work with don't generally dig into the VUS's of these reports.

And right here you see -- if I can't actually read this report anymore because I'm too old, but what this says and what my people tell me this slide says is that EGFR is amplified in this patient.

For us, in my database, EGFR, esophageal carcinoma might not be a level 1 interpretation meaning it might not be -- excuse me, standard of care of an FDA or NCCN guideline interpretation but this is absolutely a level 2 indication likely in this patient or whatever level you want to put -- level 2A, level 2B, level 3 -- but it means that there's some level of evidence that there might be a drug that might act -- and again it's not a definite.

But if you've got nothing else, this is the fourth line of therapy for this patient who has now recurred which is literally the MD Anderson cancer patient population right,you know as a clinician if you knew that this wasn't truly a VUS, a variant of unknown significance, you might start considering what therapeutic options are available for this patient.

Indeed our team, the precision oncology decision support team happened to reclassify this using our own knowledge base, we actually classified this as potentially actionable. The patient was placed on a targeted inhibitor and did respond.

This is another report that I think is very clear if you're not sure what people are doing with your patients, tests, and the patient is sent to us by their local clinician -- this happened to be a private practice clinician that sent their MET mutated patient to Dr. Hong because Dr. Hong has a MET trial at MD Anderson.

Thank goodness Dr. Hong sent this report to us, we researched where this test was performed. This test was performed in a laboratory that only performed tumor testing, did not use a accompanying germline, was reported as a somatic variant, gave Dr. Hong's trial as a recommended trial.

This variant is a known polymorphism at 48% allelic fraction. This actually -- I will give this group credit they at least give us allelic fraction so I know that they're bozo's -- excuse me I'm from AACR.

I know that their annotations are sub-optimal but at least I'd forget -- If I was from MD Anderson I would totally say they're bozos. So at

48% I kind of can interpret that a known polymorphism at half my allelic fraction is likely a genetic polymorphism in this patient.

The patients that have genetic polymorphisms at this allele do not respond to the drug so we did not recommend that trial for that patient. The patient went on to get a different agent.

Here's one of my favorites because this is where you have all these warm, fuzzy -- everybody hates me because I don't share all my data in the public domain because I'm trying to figure out how to fund a sustainable model.

And they said why don't you just put it all in the public domain -- there are a lot of these things that are crowdsourced. And crowdsourcing does have a -- can have very good value but there is a good example where a patient came in for another one of David's trials -- actually this was a dovitinib trial. The patient was recommended to David Hong specifically for this trial.

I mean the report clearly said go to MD Anderson for this dovitinib trial because you have a mutation in FGFR4 and while this is a VUS, dovitinib has been matched to FGFR4.

And I will admit my team had never matched FGFR4 to dovitinib so we thought this was a fail. I said my team failed this patient -- like we would have not annotated it this way.

So we called the company and said, "Can you just tell me because I can't find anywhere, none of these things that you've referenced actually matched dovitinib to FGFR4 mutations -- can you just tell me where that information was received?"

It was received from a public database which sounds warm and fuzzy, but it was crowdsourced data and there's absolutely no data -- zero evidence that matches dovitinib activity in FGFR4 activating mutations.

So this is a VUS. But even if FGFR4 was an activating alteration -- if this was an activating alteration, there's no evidence that dovitinib would have worked in this patient so we actually did not put this patient on that trial.

And of course, every one of these companies that provide these, give you different answers -- I’m going to be completely transparent. The MD Anderson report, in my opinion, is completely sub-optimal.

This is our report here just for full disclosure. We circle a couple of genes, I don’t do any of this work but this is what our clinicians get initially -- is they circle some genes and then they list the mutations for you but there’s no other interpretation provided.

So we get literally everything from no interpretation whatsoever to complete interpretations but that differ from one organization to another despite the fact that we’ve all talked about these consensus guidelines that suggest that we should all be adopting these things.

So while there are suggestions, we adopt them at different levels of detail, different levels of interpretation. I mean that basically every report you get is completely different regardless of what public database or private knowledgebase that they're utilizing.

The knowledgebases specifically that we looked at for ours and I wanted to just to provide these to make sure that you guys know about them but also that the VICC that Heidi just discussed, also includes these and others.

This is the My Cancer Genome knowledgebase from Vanderbilt, the personalized cancer therapy knowledgebase from MD Anderson and the OncoKB from MSKCC.

The nice bit about this is each one of these is somewhat available, so MSK for example has most of their data available but they don't have all of their descriptions publicly -- all the curated detail content publicly available.

But they are providing a lot more functionally annotated genes. So we only have 33 genes -- but all of the variants and all of the detail -- but they have 418 genes but not necessarily all of the detail so maybe if we push them all together we could get to a variant level annotation database.

But since all of us need to be sustainable the question is maybe that's something that the FDA could help try to accomplish so that we could have one standard.

Because the reality is even when you smush all of these together, this is just at the level of what we would consider and a level 1 indication almost, right? We still even differ there.

Why? Because um, things like CDK4, I'll pick on MSK, they picked CDK4 as something that has a therapeutic assertion. By My Cancer Genome and MD Anderson don't consider that a therapeutic assertion.

Why you asked? Palbociclip is an approved drug but palbociclip which would act on this pathway is not approved based on this biomarker -- it's approved for breast cancers that have specific other biomarkers but CDK4 is completely irrelevant to use of that drug in that.

So we all interpret these things slightly differently and I think that we could probably come together on some standard level of evidence that we could apply across the board -- because the reality is this matters.

I'm going to give you a case example that is a real case and a very high value patient in my life and this patient came in. This happens to be a report generated by -- we generated a report -- just so you know in the clinical environment we generated -- MD Anderson generated a report in the research environment and we generated a report at a different commercial laboratory.

All three reports, despite the fact that sequencing is terrible and it's chaos came up with the exact same 100% overlapping data, okay, so all the mutations are the same.

So my problem is not this -- my difference between what we would do and what another group would do -- again same consortium, same types of rules, we have very overlapping sets of how we would classify these things.

This group called this ALK alteration, R401Q, as a hotspot, that's what the flame means, that has some kind of -- and that gene has some kind of potential drug.

Down here -- oh I think I have this animated, and later on they indicated that there was a therapeutic match potentially indicated. Down here there's CDKN2A/2B loss and there's no match indicated because that's not at the level of evidence for that group that would merit lifting it up.

Our report actually called that exact same variant. They called it likely pathogenic basically. I called it -- my team called it likely benign. Pretty close to that same example that Heidi said right where you have everything from VUS or whatever to definitely pathogenic.

Two different groups, same -- similar sets of rules, but our group says do not act. Likely benign tells our clinicians we don't think this drug is going to work and the reason why -- just so you know why ours was so different -- is that we don't actually consider our 401Q, that specific mutation, the hotspot.

The hotspot that's found in all of these -- that's found in all the public databases is actually a nonsense codon at an R401, not a missense mutation. I consider those two distinctly different events in the genome.

The other group does not. One group would tell my clinicians ALK inhibitor trial, one group would tell my clinicians CDKN2A inhibitor trial.

I'll let you know in a couple months which one we go on and we'll give you the conclusion. I will argue that -- I'll just say that my -- our clinicians are going for the CDKN2A right now, even though that's probably a bigger risk.

This I just put out there for my own perspective but again it echoes the concept before that this one drug-one gene CDX model is no longer going to work in the age of panels and we need to be developing not only the right panels for patients but also the correct interpretations for our clinicians.

And we need to be doing this in real time so these reports need to be generated again at every time of the patient care so that as the information and as the status of the patient changes, we remind clinicians that this data is in the record and it needs to be updated accordingly because all of the underlying foundational evidence in our knowledgebases is also changing over time.

And I believe that when we do this correctly -- so this is from unpublished data that we're trying to get published, so if we could find somebody to accept it this is the line of patients who were tested -- all of these patients were tested on the same assay, all of these patients were found to have a mutation in an actionable gene, that we would have considered actionable.

These patients did not match -- were not matched to any drug. These patients were matched to a drug and there is a statistically significant survival impact simply again just based on matching to agents.

So we believe that if we do these right and we give the physicians the correct interpretations, whether you're using a public or a private knowledge base we can improve patient outcomes, and here are the acknowledgements.

DR. MADISON: Thank you Dr. Shaw for that talk and giving us an exciting perspective on what you guys have going on in the AACR GENIE and how we can utilize the databases effectively.

Our next speaker is Dr. Ben Park. He is a Professor of Oncology in the Breast and Ovarian Cancer Program at the Sydney Kimmel Comprehensive Cancer Center at Johns Hopkins University and a Physician Scientist with a focus on exploiting genetic alterations for diagnostic and therapeutic purposes.

He is also Associate Director for Education and Research Training for the Cancer Center and Associate Dean for Post-Doctoral Affairs for the School of Medicine, Dr. Park?

DR. PARK: You've probably noticed there's a lot of overlap. Kenna and I know each other from the AACR GENIE Project so this is going to be a little bit redundant.

I was going to start off by saying now for something completely similar. But hopefully, you'll also see that there are challenges that have been brought about and I think a lot of what we are doing right now are kind of repeating what we do for specific tumor boards anyways, that there is going to be difference of opinions between one institution and another.

I will tell you I think our tumor board is more in line of what Ken was saying -- we really dig into the weeds and look for data like, is this a ? mutation if it's a hotspot -- a reported hotspot mutation, et cetera.

So these are my disclosures. I'm going to start -- again we've heard a lot of this before but the difference between how we do germline testing in this country versus somatic testing or tumor testing I should say.

And the classic example I think everyone knows is Myriad's genetics testing for BRCA1 and 2 -- that's been around for decades now literally. And obviously testing has evolved with NextGen sequencing.

As we've just heard there's panel gene testing which has become the norm. Typically, relatively still a small number of genes but now this is actually expanded from 1 to 90 up to 30 to 40 and with that has come some complications.

Testing for germline as again was iterated in the past has always been -- or not always, but we've always recommended should be done in the context of genetic counseling.

And usually patients who are deemed higher risk are the ones who actually go on to get testing. That's currently an evolution. I think there's been a lot of provocative data that metastatic prostate cancer patients may actually have a higher rate of germline mutations and DNA repair genes and I think there are a lot of similar studies going on right now.

Important though as this last bullet point says this requires consent ahead of time and we've already heard about the challenges of that if you're dealing with paired sequencing of germline with tumor.

In contrast tumor only testing just to again reiterate in the past was really very, very limited to just hotspot mutations and genes and then the larger cancer gene panels come up for commercial as well as academic and we now know that whole exome sequencing ? probably is going to be on the horizon because this technology keeps getting faster, better, cheaper.

But whether or not that's the right thing to do and how we are going to interpret the data which is already complex is going to present itself with huge challenges.

And I think the second bullet point is something that we are struggling with as well as everyone else. How do you really distinguish the true somatic alterations -- that's very problematic. We generally do not require consent though this is an evolution – is a --changing as you all heard.

And although there is utility for some mutations, when people have tried to use the traditional benchmarks that we do in clinical medicine -- that is if you take the whole group and sequence them, what is the kind of overall benefi?

I don’t think that that’s right now because as we heard there’s a lack of drugs -- is rarely going to ever be a positive study. But I do believe that for those select patients who you have a really bona fide, druggable or targetable mutation with the right drug you can actually do a lot of good.

And for the most part, at least at our center, this is actually done in metastatic disease. There’s really still, I think, very little clinical usefulness or utility for early stage solid tumors.

So we’ve heard a lot about this but I’m going to go through this again and present some examples because there’s even more layers of complexity and subtleness that you will find and it can leave you scratching your head.

So we sequence both normal germline DNA along with a tumor tissue -- this allows for the filtering of germline variants and when you’re doing bigger and bigger panels and eventually whole exome sequencing this becomes almost essential so that you can filter out the noise from the signal.

So the true advantage obviously is that you can see just what's in the tumor. The disadvantage though is that you're potentially filtering out really important things in the germline that in the past were more about relative risks, inheritable pre-dispositions, but now we also have drugs for certain germline inheritable mutation, BRCA1 and 2 with PARP inhibitors and MLH1 and MSH2 and other mismatch repair genes -- now we have ? checkpoint inhibitor therapies.

And so these patients are not consented for germline testing. Companies and academic centers can't really report them but they have now embraced the idea of being quasi ambiguous in saying in the appropriate genetic context, in the clinical situation, they would recommend further germline testing.

So some of these tumor-only tests state that -

- where it’s become even more confusing is because of this recognition that you may be missing something -- some companies are actually doing a little bit of a hybrid model.

And if you’re not knowing exactly what is being sequenced and what is being filtered as some of my colleagues have already mentioned, you could really get led astray.

So here’s an example, this is a company that has now decided for whatever reason just for these four genes that they will not do electronic filtering so they do perform germline analysis and they use it to filter out tumor, but if you don’t go to page -- I think Ken had said 23 -- I don’t know what it is for this particular company, but if you don’t go to page “x” which is in the back of the report and recognize that they’re not filtering out these four genes, you may actually think oh, everything that’s being reported here is only somatic and that might actually not show up anymore or would show up.

This is again the layer of complexity that’s going on with the industry right now and so knowing exactly what's being tested is incredibly important and how that affects your recommendations for the patient both for germline testing as well as therapy becomes incredibly important.

So as this slide states, one needs to understand what is being tested and what is not being tested. Here's another set of examples -- this is from a company that did tumor testing and as you can see from the slide there they present not only the purity of the tumor specimen, so in this case 75% tumor purity.

They also present what is the source of DNA -- in this case it's saliva. Now the same company will get another sample and if for whatever reason a source of normal genomic DNA is not presented, that's what you get -- in small letters -- this is a little bit blurry and I was going to retake it but then I thought no, it's good that it's out of focus because it brings it to the point that this is a little blurry.

And so the source of normal DNA here says, “not provided,” and the mutational report from this particular prostate cancer patient -- so you look at that antigen receptor gene and there’s a missense mutation.

If you look at that missense and with the tumor purity of 20% but a mutation allele fraction of 100%, most of us in this business would start scratching our head and thinking that doesn’t sound right.

And so sure enough if you go into the public databases and look at all the different things like ClinVar that we had, this has in fact been reported once as something that could lead to antigen sensitization as an inheritable kind of a gene variant but it is also reported as having no effect.

And so really to me this is still a VUS, this probably had no bearing I think on this patient’s development of prostate cancer. Whether it had any bearing on responsive therapy we don’t honestly know, but certainly this is probably going to be germline.

Whether or not it merits getting germline testing is a whole other question because again this for me is more like a VUS. And when you look across again the different genes and the different types of allelic frequencies the thing that really probably sticks out here is the PIK3CA mutation.

Again we don't have great -- well there is now one PIKinase inhibitor approved in a different cancer type but we don't have the definitive data yet for prostate cancer and PI3Kinase inhibitors.

This would be someone though that we know that that mutation is an activating mutation and if they could get on a clinical trial, that that would actually make sense.

The other thing that a lot of companies are reporting -- and I don't know why they are still doing this because we have data to suggest otherwise -- that PIK3CA mutations are going to predict for response to mTOR inhibitors.

And even though a lot of clinical or pre-clinical data I should say speaks to that, including some of our own work, somewhat embarrassingly now, this is not the case.

We know from various clinical trials that PIK3CA mutations do not have predictive ability for MTOR inhibitors and yet that's still in the majority of reports that are out there.

And then there are always caveats. This is a particularly interesting click case that also got into this whole realm of what we call clonal hematopoiesis but this is something actually of that on steroids so to speak.

So we have this very interesting case in our tumor board where someone had a duodenal tumor and that tumor was resected, it was a metastatic patient so the testing was sent off for a NextGen sequencing company and it came back without any allelic fractions that said, "Oh, there is this JAK2 mutation, V617F."

This is kind of the driver mutation for a blood disorder called polycythemia vera, and there's a drug for it which again some of my predecessor colleagues have already mentioned --

ruxolitinib and we thought wow, this is incredible let's give this patient ruxolitinib and see what happens.

We discussed this at our molecular tumor board, we had our hematology oncologist or heme malignancies colleagues there as well and then it occurred to -- well something happened, I went back to the referring physician and then he said, "Wow, that's really interesting Ben," because she has a history of polycythemia vera.

And I had one of those "ro ro rastro" moments if you know what I mean is that oh God, what is going on here. So we actually ended up repeating this and you can see from the slide there we did our own internal NGS on the duodenal cancer and you can see the allelic fraction is only 13%.

If you look at the tumor where the arrows are those are kind of pools of blood that were in fact contaminated in this tumor tissue and the adjacent normal tissue was actually also positive at a low, low, allelic frequency.

We wanted to be really sure that in fact this was, again, not in the tumor but also not in the germline. There have been some rare familial disorders of not necessarily polycythemia vera but a related disorder of having this very mutation in the germline.

So we did a buccal swab and we just Sanger sequenced -- we didn't want to get to a low level of allelic fractions. Weirdly it looked like it was germline. It was 50/50 and then our genetic counselor who sits on our tumor board basically said then you can't use cheek swabs because they're going to have lots of polys and in fact that was the case.

So we actually did a fingernail clipping using a forensic pathology kit. I'd like to say we nailed this one and that was completely well typed.

Alright, so this is our tumor board, we call it genetic alterations in tumors with actionable yields or a gateway. We have published on this and our kind of actionable mutation frequencies is also about like 10 to 13%.

So again, I think that's becoming a little bit standard despite having some heterogeneity in opinions.

This is our purpose -- I kind of already went over that I'm not going to go over all of this, as well as our mission statement which again is pretty self-evident after this whole session.

Our definition of actionable is very akin to other people. Again, does it have an FDA approved therapy in the right cancer type, in a different cancer type, something that actually provides rationale for a trial, but importantly also genetic alterations in the germline and what are the consequences of that?

We have additional considerations that I don't think we have discussed here yet but things like should we be giving a targeted therapy now versus standard of care therapy.

This is especially relevant in my disease, breast cancer. I already mentioned about potential germline variants. One of the things that has come up in the Michigan's sequencing effort that was published -- or not published, but reported in the New York Times -- is what do you do with this example where they actually found integrated pleural HIV DNA in the cancer specimen because they're doing whole genome sequencing at the time?

What are the ethical and legal implications of that and I think that's why we've also added ad hoc legal input as well as emphasis.

And then liquid biopsies which are near and dear to my heart. So for cell free DNA most of these tests do display allelic frequencies. They're again germline variants and clinical hematopoiesis which I kind of mentioned which are just the beginnings of myeloproliferate disorders or myelodysplastic syndromes and they're usually actually pretty easy to spot if you look at the type of gene mutation and allelic frequencies.

But again there can be confusion. So this was a 73 year old with metastatic breast cancer that was originally diagnoses in 1998 and she had a strong family history.

She never actually got germline testing but she did not have an easily biopsy of the lesion. So a cell free DNA test was sent off -- she had ? mutations which made sense but if you look at the BRCA1 with the asterisks, that's a stop mutation and you might think that that's real but the clonal frequency or the real frequency made it a sub-clonal so you weren't really sure.

And you don't see an allelic frequency of 50% that you would think would be the other kind of inheritable allele. So she did go on to get germline sequencing and it was in fact normal.

And one of the things that you have to think about though even if you don't see that, there are rare examples and I have a couple of patients like this where the whole gene is actually deleted in the germline and so a lot of these blood tests are not going to be set up to detect a single gene deletion relative to the wild type allele.

I think I pretty much went through that -- so what we decided that initiating a PARP inhibitor right now would not have meaningful benefit so instead we actually recommended a PI3Kinase inhibitor trial.

So, in conclusion I think one really has to know and understand not only subtle -- or differences between tumor testing and germline testing but the subtleties that companies may or may not do in terms of what gets filtered out and what doesn't.

You have to have great care when interpreting these tests and these results, knowing what is being tested and what is not is of paramount importance and recognizing caveats that we and many others are actually discovering.

Keeping up to date with the literature and clinical trials is extremely difficult because this is a fast moving industry but it is absolutely necessary if we are trying to do the best for our patients.

And things that we are recommending today versus four years ago have changed. And I think establishing the molecular tumor boards can really help with that.

You know one of the things that Kenna brought up was that EGFR amplification. The reason Foundation actually puts that in the VUS is because when they get cases like that -- we ran into a similar thing -- whole regions of that chromosome where that gene is located are amplified.

And so the reason that it's done, I think, that they just stick it there without an explanation is that they're not sure whether that's truly a driver for that cancer or if it's a passenger so to speak because it's just been co-amplified with multiple other things.

So those are again the layers of complexity as well as subtlety I think everyone who's interpreting these tests really should be aware of and here at Hopkins or a little bit north we're actually trying to not only do our molecular tumor board, but we're also setting up kind of courses to really help community oncologists, other academical centers set up their own tumor boards, and that's it, thank you.

DR. MADISON: Thank you Dr. Park for that great talk. We have one last speaker for this session. We have -- next up is Dr. Karla Bowles. She received her PhD and completed an ABMGG fellowship in clinical molecular genetics at Baylor College of Medicine.

In 2006 Dr. Bowles joined the dedicated professionals at Myriad Genetics where she is currently a Senior Laboratory Director and serves as a director lead on the variant classification team. Thank you.

DR. BOWLES: So I'd like to start today by thanking the FDA for allowing me to come and speak with you and I would especially like to thank Dr. Madison for all of the work that she put into organizing this particular session.

As Dr. Madison said I am employed by Myriad Genetic Laboratories and I do receive salary and stock options as compensation. I want to just sort of begin today by saying -- very similar to Dr. Rehm's talk -- the talk that I'm giving today is really going to focus more on hereditary cancer testing and germline testing although you can see that there will be some implications to somatic tumor testing.

When we consider a variant classification, variant databases and their associated tools, we really need to consider variant classification and re-classification at the same time.

While they are two separate processes, they are still very closely related to each other. When we initially observe a variant I think most laboratories, whether they're academic or commercial laboratories, attempt to classify that variant in terms of a five tier classification system which is supported by the ACMG guidelines.

The classifications of pathogenic and benign are considered to be definitive classifications in that they have reached their endpoints as far as variant classification is concerned.

However, when we think about variants of uncertain significance, that is definitely not a definitive classification -- and even likely pathogenic and likely benign classifications, those variants still have an additional step that they can go before they reach their endpoints.

And so we hope one day to re-classify those variants as more data is gathered and we deem that data to be sufficient -- hopefully we can move those variants into a pathogenic or benign definitive classification category.

At Myriad when we were developing our database we had several clinical key questions that we had to ask. What data quality and accuracy standards will we require? How will we maintain and document that integrity? And finally, how often will we update our database and variant classifications to serve the needs of our patient population?

I'll begin by addressing the first question and when we consider this question regarding quality and accuracy standards, we really need to understand that this is a balance between classification speed -- how fast are we going to make it to that definitive classification -- and classification accuracy.

We can get there fast but are we going to be correct in the final classification that we assign? We really need as laboratories to decide for ourselves what classification accuracy thresholds we are going to mandate.

At Myriad over the years we have examined multiple data quality options and we have determined that we must base variant classifications on high-quality data due to the often irreversible clinical implications associated with positive and negative test results.

So while we've examined multiple database structure options we have ultimately chosen to go with option number 1. Our classifications are based on very strong and strong data as defined by the ACMG classification guidelines.

And importantly, we set high accuracy thresholds for all of the internal classification tools that we use. So when we use an internal classification tool to classify a variant as likely pathogenic or likely benign, we require greater than 99% accuracy for that tool.

If we're going to go all the way to pathogenic or benign we require much greater than 99% accuracy for the tools that we use. While we could go with options number 2 or 3, it's important to understand that high quality data is very slow to obtain.

It's much easier to obtain lower quality data which will drive us to a definitive classification quicker, but by lowering our accuracy thresholds, our data quality thresholds, we will be introducing significant errors into our variant classification database and those errors will ultimately end up on our patient reports.

One of the key factors in establishing a high quality database is to establish classification confidence thresholds before we use data.

When we examine our internal tools at Myriad we estimate accuracy for each of our classification tools independently.

Each tool is evaluated independently using large numbers of control variants. We also estimate tool accuracy for each gene. Over the years we have learned that some classification tools may be great for some genes but they don't work quite so well for other genes and it's not always safe to assume that the accuracy of a particular tool is uniform for all genes.

Tool accuracy is also estimated based on clinical effect. There are multiple examples of many variants in the scientific literature and laboratory practice where we can find a variant that has a significant functional effect on a protein or a protein production effect, yet that effect does not seem to quite translate to clinical effect or high cancer risk.

We believe that there should be a more direct connection between the classification tool used and the actual risk of cancer. Because of this we often exclude tools that are based on lower model organisms.

So for example, if you can imagine trying to translate a yeast protein functional defect and a human cancer risk -- that's really quite a distance for any tool to have to go and because of that we often exclude these model organisms.

Finally we used unbiased tools whenever possible. High quality statistical tools have a quantifiable accuracy. In contrast, tools requiring significant human interpretation have a much greater chance of error.

I'd just like to show you one example of how we would evaluate an internal tool. Several years ago Myriad published pheno analysis under the name History Weighting Algorithm and the two references that you could see at the bottom left of the slide.

And this is one of our primary variant reclassification tools. Pheno is a statistical tool that classifies variants as pathogenic or benign based on whether or not they're associated with strong personal and family history of cancer.

Pheno is highly accurate and we have developed and validated it for each gene for which we use it independently. We did this using large numbers of positive and negative control variants of known classification, between 32,000 and 79,000 variants depending on the gene.

Based on this analysis we can determine that Pheno has positive and negative predictive values of greater than 99.5%. Importantly Pheno measures the association of a variant with cancer risk, not the functional effect of the protein.

So once again we have that more direct line as to whether or not this is associated with increased cancer risk. And finally Pheno analysis is one of those statistical tools that does not rely on human interpretation.

If we give the Pheno the same data over and over again, we would expect Pheno to always come up with the same classification. In contrast subjective classification tools rely heavily on human interpretation should always be used with caution.

On this slide you can see three examples of subjective tools -- literature review, the analysis of population data and structural analysis.

While all of these are very valid, heavily used variant reclassification tools, you can see that there are quite a few questions surrounding each tool which really requires that they be addressed with human experts.

And that leads us to the next question in terms of a variant classification database. How will we maintain and document database integrity? One of the ways that we addressed that at Myriad is to have a classification committee of experts who maintain classification accuracy.

We determined a long time ago that just having one or two individuals reviewing and classifying each variant is most likely insufficient for a highly accurate database.

So we have a variant classification committee composed of our laboratory directors, genetic counselors, PhD level scientists who are experts in their fields as well as variant specialists and this group meets on a daily basis to review all of the novel variants that have been seen at Myriad within the last 24 hours and reclassify those variants as a team with each person weighing in from their particular area of expertise.

Our classification process leverages the human strengths of our variant classification team as well as computer automation. Our classification process uses a combination of manual and computer-assisted and computer automated steps to analyze approximately 50 to 100 novel variants per day.

The first step of this process includes automated analysis of each variant by a Myriad developed computer program and database called VITA.

VITA helps us analyze each variant on a variant by variant basis. It starts by gathering variant specific information such as functional domains, gene locations and many other parameters and it makes -- puts this together to present to committee members.

VITA, based on the data that it gathers, along with pre-defined SOP requirements, proposes an initial classification for each variant.

However, its expert committee review that assigns the final classification based on the data generated and curated by VITA, data from peer reviewed literature, data from internal tools and potentially data from other sources if that data becomes available.

One of the key aspects of our variant database VITA is that it used a queue based system to enforce appropriate human review. In this particular example you can see an MSH6 variant which will pass through all of the queues that you can see in the orange box.

At the beginning of the process VITA will pass that variant to two variant data specialists who will perform preliminary analyses. After those analyses are complete, the variant will be passed on to a series of queues -- not all which are shown on this slide -- where PhD subject matter experts will perform more in depth analyses.

After those analyses are completed, the variant will be passed to a new mutations committee who will do a thorough review and assign a classification to that variant based on their analysis.

And finally, VITA will pass that variant to two laboratory directors, the first who will write the report text and enter the final classification into our database and the second director who will confirm that classification and report text.

And this way we're assured that all analyses have been performed. We believe that a well-controlled variant database is critical for quality. As I said our database uses a queue-based system to enforce appropriate human review.

It will not allow a variant to be classified until all reviews are complete. Our database also enforces the classification of the variant itself. It requires verification of a classification by multiple individuals.

The database alerts users to unexpected classifications and a final classification by the laboratory director must agree with the committee decision or VITA will not allow the classification to be saved.

Finally, VITA provides an audit trail for our database. We can see who was involved in a particular classification, when the variant was classified or reclassified and what specific data was used in the classification of that variant.

Despite all of the wonderful tools that we have at Myriad we still find that VUS are unavoidable and so that brings us to our third question -- how often will we update our database and variant classifications in order to meet the needs of our patients?

There are multiple approaches that can be taken to this question and this slide shows some commonly used approaches. The first is to review and attempt to reclassify each VUS on an annual or semi-annual schedule.

Another approach would be to review and attempt to reclassify each VUS every time it is seen in a new patient. In some cases that means there might be a few weeks or a few months in between reviews, however, if a variant is rare it may be multiple years between reviews.

And finally, the third option -- which is the most labor intensive -- is to implement a near real time review process. When Myriad considered options number 1 and 2 we could quickly see a very large pitfall.

Imagine a hypothetical variant shown the left here and we'll just call it BRCA1 variant B which is initially seen in a patient in September of this year and is classified as a VUS. Shortly thereafter a paper is published which definitively shows that it's pathologic and somewhere down the road that variant comes up for annual review.

It may be a year or more between the time the data was available to call it pathogenic versus the time it's actually upgraded to pathogenic. During this time, the patients initially identified with that variant will have a VUS report in their hand and be clinically managed on that report, but they would more appropriately be clinically managed as a pathogenic mutation carrier.

And this is a time period that's really a missed opportunity for those patients. It's a time period when a cancer could have been avoided if they had pursued prophylactic surgeries.

It's a time when a cancer may have been detected at an earlier and more treatable stage, and it's a time that family members may have benefitted from genetic testing.

Therefore, Myriad has chosen to pursue and implement a near real time variant review process for our patients. One example of the way that we do this is through real time evaluation of the scientific literature.

Before we launch a test at Myriad we do a complete literature search to identify all variants previously published and we upload that information into our database along with their associated papers.

We perform a daily literature search of all the literature published within the last business day where we continue to update our literature database.

We do another search on a variant on its first observation to make sure that we have captured all of the relevant papers and then we continue to do daily monitoring, even after a variant is re-classified.

Another way that we keep up on our variants in near real time is through the automation of our statistical tools and other classification tools that we use at Myriad.

For example, if we go back to variant B which was classified as a VUS -- in addition to keeping up on that literature in real time, every time we receive a new sample from a patient and we identify that variant our statistical new tools, which run on the background of our computer 24 hours a day, 7 days a week, will reevaluate the data from that variant.

And if we now reach a statistical threshold, the computer will email our new mutations committee and let us know that we have a variant that we can reclassify.

We will bring that variant to committee

review typically within one business day of the new patient data. We'll get that variant reclassified and we'll send amended reports, roughly within 20 -- within 7 days.

We believe that a robust variant reclassification program is in the best interest of patient care. When we look back on 2016, Myriad alone reclassified 529 variants based on our automated tools and our variant reclassification program.

That's allowed us to send out over 23,000 amended patient reports to individuals who now receive a more definitive variant classification.

Future patients will also benefit from this as they will receive definitive classifications rather than uncertain test results. And finally, we can never forget that this affects not only our patients but also all of their family members for future generations.

So in summary, when we look at our approach to a variant classification database -- a clinical database, we believe that data must be of high quality. We have established high variant classification thresholds.

We use unbiased statistical tools whenever possible and we have an expert variant classification committee to insure its consistency.

We believe that database integrity must be maintained. Our database has full traceability. We can say who, when and what specific data was used to classify or reclassify a variant.

And finally we believe that our database must support ongoing variant monitoring and reclassification and the issuance of amended reports.

We've developed innovative classification tools -- we perform near real time monitoring of the scientific literature. We've automated our statistical analyses and we've set up a robust program for the notification of healthcare providers regarding variant reclassifications through our amended patient report process.

And with that I would like to thank you for your time and turn the platform back to Dr. Madison.

DR. MADISON: Thank you Dr. Bowles for that talk. I would like to invite all the speakers for Session 3 up here for the panel discussion. I want to give a round of applause to all of our speakers for this session.

And so similar to the previous sessions we'll have some moderated Q and A here and then we'll also open it up for public question and answer.

So first thank you all for your wonderful talks and I think I want to start with you, you had a good range of information here provided so some on the database side, some on the using of that information for clinical interpretation and some of the nuances and the caveats associated with giving this information to patients and the clinicians to use correctly.

I want to start with the database questions. So one of the things that was noted -- that the information attached to the clinical assertions in the databases can range from no information at all to very detailed curation.

I wanted to get you all's perspective on what do you think is the necessary level of metadata that should be attached to the variant assertions that are provided in public or, you know, private utilized databases.

DR. REHM: If I can start. In my view the transparent rationale for how that variant was classified needs to be provided and that's what most of the submitters to ClinVar do, but particularly, some of the older submissions don't have that, some of the labs that just don't have that data separated from their patient data ,haven't yet constructed that submission.

But I think the evidence that formed the basis for your classification needs to be there or linked in some way.

DR. SHAW: Yeah, I would agree with that also. I think that often we -- even without the exact rules, as long as the databases provide us with the Pubmed IDs, or the abstracts, or whatever they're using for their evidence that they used --

we can go back to that and determine whether we would have applied the same rules to that interpretation.

So we need to know the evidence that was used to provide that -- that classification.

DR. PARK: I really don't have much to add except that I agree with that. I think the transparency issue is often one that is not there -- meaning you have no idea. There are companies that will do the annotation for a lot of the companies that do the sequencings.

And many times you just have no idea what their algorithm is and how they pull out the data.

DR. BOWLES: And I would concur with everybody else. There really needs to be enough data attached to each variant that you can look at the variant and very clearly see what the rationale was to make sure that there's enough data there that you can either agree or disagree or at least know that you need to have an intelligent conversation.

DR. MADISON: And so one of the things that was highlighted as well -- when you have these databases that the rules that are applied are sometimes very clear and very understandable but then there are other times where you have no idea what quality control measures are set in place to incorporate the rules for it then, you know, once the database information is provided and then reported out to those who are going to be users.

I wanted to get your ideas and your thoughts on maybe are there specific QC measures that are necessary for say -- somatic databases, does that differ between germline databases?

And if you have any thoughts about how you measure the validity once you see those rules and whether those are valid rules to place this information within the databases?

DR. REHM: So I guess I sort of see it in two different ways. One is the rules that the individual or laboratory or source classifying that variant use to call it what they called it.

Are they using the ACMG guidelines, the AMP guidelines, you know, or their own custom approaches to variant classification? So like for example in the ClinVar database, anyone who is a single star submitter has to submit their rules and methods for variant classification if they're at that level -- the same with anything above it -- expert panels, et cetera.

But then a separate question is the owner of that database which could be one lab or in the case of most public databases is another entity. You know, are there algorithms being used by that entity that's overarching?

So for example in ClinVar there's algorithms for how it takes many submissions and gives an overall clinical significance and it's based on hierarchy of three star overrides, you know, lower things.

So I don't know if you're -- and that obviously has to be very transparent but also you could always go down and see what every individual lab said in case you want to see the granularly that went up to it.

So as long -- so I think the take home

message is you have to be very transparent about the rules either for an individual classification or for an aggregation of information as well.

DR. MADISON: Good point.

DR. BOWLES: I think importantly when we look at the rules -- um, you know, what data does it take to classify a variant as pathogenic or benign? I think every database has to be held up to the highest standard -- just submission to the database itself, even with a short explanation is not necessarily sufficient.

As a laboratory director it's all of our responsibilities as far as the accuracy of each variant classification that leaves our particular laboratory.

And so we need to really have solid access to the primary data. If -- when I review a functional assay in the peer review literature, I require that that assay meet certain criteria.

We require that same criteria of a functional assay that was maybe done by a research or even a diagnostic laboratory and then cited in a database

-- all of that information has to become available.

We can't just take the word of whoever entered into the database if those experiments were performed correctly.

DR. SHAW: And I think the other issue is that even when the rules are transparent you have to determine what your level of risk is. So I'll give an example. So for us we, for example, if there's a mutation -- I'm just going to say V600E -- the next patient, obviously I think we all agree on that, but the next patient comes in with a V600L -- whatever.

It's never been seen before. We would not classify that as actionable. Other variants -- other databases actually do because it's a variant at that location and it's been seen to be actionable at that location with a different variation before.

Other things are what we do is P10 -- if it's truncated -- because the truncation can happen almost anywhere up in the early part of the protein and cause a loss of function we think.

We would consider that potentially actionable even if that variant had never been seen before if it loses all its functional domains -- that might be more aggressive than you would want to apply for your own database.

I think you have to really understand what rules are being applied and how aggressive or conservative you want to be in terms of how you use those for your own patients.

And so it's not enough just to understand the rules but to determine at an institutional perspective where your risk tolerance is for maybe getting it wrong -- because this is still not an exact science for a lot of these variants.

DR. PARK: I would just amplify on that too that part of what we're trying to do is when we look at the levels of evidence we think about it both in terms of pre-clinical, some of what was being talking about in the, you know, laboratory.

But for most of us clinicians that's really not enough. That's enough to maybe say you could get on this trial but we would certainly never recommend -- at least at Hopkins, off label therapy for that.

On the other hand, if there's some evidence out there in a clinical setting, whether it's an N of 1 or a case series or even a full-blown trial and depending on the nature of the trial and the results, we might be a little more comfortable recommending an off-label use.

That's where I think one has to also sit down with as far as QC -- what are the studies one is looking at? Preclinical or clinical -- and then if it's clinical what are the different tiers of evidence that we can derive from that?

DR. MADISON: You actually you know bring me to my next question as you noted about the various studies and one of the things that Dr. Rehm noted within her presentation was that a lot of the clinically significant conflicts were really based on literature-only sources or largely based on literature -- what do you do with that?

Like how do you utilize that information when initially you always think of literature as a really good starting source for getting, you know -- that type of background information for your evidence?

DR. REHM: Yeah I think you know, one of the challenges with the literature is um, it's mostly quickly outdated and largely not updated over time because most research studies are um, a point in time. They aggregate everything they can, they publish a ? paper and then they move on to the next study and so that information gets out of date and it doesn't get maintained.

And so in the end there may not be a real discrepancy it's simply that that's an out of date interpretation and ClinGen is actually working on a project to represent the ClinVar data and remove some of these out of date older things where they just didn't have the same evidence at the time they made that classification.

So taking into account the date an assertion was made is really critical. The other thing that I think is a challenge in the public literature is there's a bias in terms of publications and wanting to make your story seem more interesting and so there's a tendency to over-interpret I think, in the public literature, whereas in a clinical lab the end of the day liability sometimes actually goes the other direction -- so there's forces in either direction.

But I think um, the literature is particularly susceptible to that desire to over interpret. And sometimes it's just that the variant was through into a table of all variants seen in patients with disease and there's this implication that everything is pathogenic but the authors actually didn't state that but they get dumped into -- like the HGMD database.

Anything in that table just gets labeled DM, you know deleterious mutation. So I think there's this challenge of well what was really stated and documented in that paper versus what wasn't and then the bias in over interpretation that we always see.

DR. MADISON: Thank you.

DR. BOWLES: I think another thing that we do within our practices when we try and review the peer reviewed literature when we look at different functional assays we have to assess not only is that assay applicable to cancer, but to try and put an accuracy estimate on any particular assay -- has it been performed and replicated in multiple laboratories?

Has it been performed using enough variants of known classification that we can determine whether that assay is accurate 99% of the time or is it only accurate 80% of the time which may be sufficient for a research study, but it's not sufficient for a clinical test.

And if there's a quantitative aspect to that assay what is the correct cut-off versus what is the cut-off that the author proposed.

DR. REHM: Yeah I just want to emphasize Karla's absolutely right. What we find when we are doing variant discrepancy resolution, the largest source of discrepancy between laboratories interpreting variant is the subject of interpretation of functional data degree in the publications.

And one person looks at a graph that says that there's an effect and says, "Well that's good evidence." And the next person looks at it and goes, "No, that wasn't well validated -- they didn't validate the assay with known, you know, pathogenic and benign variants and so on."

And this is where we really tried to bring on our ClinGen expert panels -- people who really understand these assays, can determine how well they're validated, replicated, what the quantitative cut-offs are and guide the community in how to use these types of assays, because a huge percentage of them really just are not well validated and that's an important point.

DR. MADISON: So you've led into my question as if you knew it before. One of the things that Dr. Park noted in his talk was some of the nuances in receiving outside data and really digging down and interpreting what that really means and whether or not something that says it may be actionable by a clinical lab, that a result that your patient may have -- has received, but the information that you guys have gathered presents something different.

And I want to note the caveat that you all have noted multiple versions of a board -- a panel of experts, a group of people who review all this information and really dig down deep and get a better understanding of what it truly means.

But when you think about some clinicians or hospitals who may not have that level of access or expertise or are able to you know, mine through data, or really, truly understand some of the underpinnings -- how are they able -- what would be your tips to help them understand how they can ensure that the diagnostic data that they are getting is accurate and reliable?

And then, the clinical databases or the public databases that they are going to, to put that information in and try and get and get some interpretation out -- is accurate and reliable?

DR. PARK: I was going to say that I think that is a huge challenge for us right now as we've grown our tumor board we're getting many more requests from community physicians and others.

And so my kind of take on this is that beyond providing this service we actually have to be educators. And one of the things we're rolling out right now is training other people to eventually be able to run their own molecular tumor boards so we are actually consultants for something called the Maine Cancer Genomics Initiative.

We've actually been helping out with other tumor boards locally and nationally. We're working with Allegheny Health Network, et cetera.

And we dial in, teleconference, et cetera, but I think at some level -- you know this is like any other type of process. The more you do it or the better you get at it and the more comfortable you are how to actually do this to yourself.

So I'm a big believer of the, you know, teach a person or give a person a fish and he eats for a day or she eats for a day and teach them and they'll eat for a lifetime, because I really don't think any one academic institution given the amount of testing that's now going on is going to be able to sustain and have the bandwidth to do everything.

DR. REHM: And I will add that you know there's no capacity for all of us to follow-up on every individual piece of data, whether it's in the literature, it's in the database or you just want to track down what you might find.

But when you are in that situation where there's a variant -- and I forget who made this comment, the vat or the variant of almost significance -- somebody said that earlier today.

You know, where you think that's going to make a difference in a patient and you're really looking to make, you know, one additional piece of data or there's some discrepancy that looks a little fishy and you want to dig in and see if there's a miscommunication or something -- that's when a lot of us, you know, go that extra mile.

You know I'll get a rec form that says this patient’s affected but the variant doesn’t segregate um, and that doesn’t make sense. So I call up the physician and say, “Gee you checked off this family matter as affected, can you describe to me the actual data for that individual?”

And you know, let’s say you’re in a hypertrophic cardiomyopathy case -- they say well I checked it because the patient fainted once. I’m like well that’s not a diagnosis of a hypertrophic cardiomyopathy -- people faint all the time.

And so in the end the variant -- that one person who looked like a non-segregation was in fact, misdiagnosed and that check box was not adequately checked off.

So those are the kinds of things that we all do to follow-up and you know what triggers you to do that, well you know, something looks suspicious or you’re basing something based on a -- you know, a piece of paper that someone checked off.

I mean those -- or you’re reading a

publication and a lot of data is in there and something doesn't make sense and you contact the authors and then they tell you -- and this has happened many times -- oh yeah, that grad student who put all that together, you know, the database was a mess.

They'll put that -- so you know, but you can't follow-up on everything but you do have to use your best judgment to decide when it's going to make a difference for the patient and you should go that extra mile.

DR. BOWLES: I think ultimately we need to remember that it is the laboratory and the directors of the laboratory that are responsible for the final classifications and interpretations of the variants that go out the door.

And whether that's them hiring the appropriate people and getting them trained appropriately or working with a collaborative group, ultimately it is the diagnostic laboratory that is responsible for the accuracy of the interpretations of the variants that they report and they all need to find a mechanism -- whether they're a large laboratory or small laboratory -- to provide those accurate interpretations.

DR. MADISON: Well that actually leads into my last question before we open it up to the public group here to ask questions -- is when, you know, this field is constantly moving forward very quickly.

The level of evidence for certain variants is changing day to day. And Karla you noted that there were -- you guys check the literature every day. You have a system automated to look through that.

And while that may not be available for everyone, what is the responsibility of either the clinical labs or the healthcare providers in continuing to dig through and get the most updated information and the timeline or to inform the patient -- here's some changes -- or the clinician, here's some changes that could affect how you may consider treating this patient?

DR. BOWLES: I believe that ultimately that responsibility falls to the clinical laboratory. I don't think it's reasonable or would fit within a particular -- most providers' patient practices for them to be expected to go into public databases and update the classifications for all of their patients that they're seeing on an ongoing basis.

It's oftentimes, out of the scope of expertise for many of those healthcare providers. They're relying on us as a diagnostic laboratory to interpret those variants for them and it would be extremely difficult for, you know for example -- a primary care healthcare provider depending on what they're offering screening for, to be checking cystic fibrosis, you know -- carrier status one day and then to another day have to go looking up breast cancer.

Now my next patient's in and I have to go look at the colon cancer genes. I think it's really not reasonable to ask that of physicians. It really falls to the laboratories to update their databases, reclassify the variants and send those amended reports in a timely manner.

Because often times you know, when we're talking about cancer -- there really is a window of opportunity that could be missed for our patients.

You know, as I said if you're only updating a variant once a year or once every few years, those really are those missed opportunities. You could have caught a cancer before it happened.

You could have given more aggressive surveillance and caught it at an earlier stage. And we always have to remember in my world -- which is the hereditary cancer world, this affects generations to come.

So even if you get 10 years down the road, 15 years down the road, it still matters to that patient and it still matters to that patient's family members.

DR. SHAW: At least in the somatic space I think we have it slightly different because I think that for us it's more at point of care is when it’s important and so we’re trying to provide -- at least at our institution is the ability -- we have two things.

One, if we do reclassify a variant we don’t do it on a continual basis every night bringing in new data -- although that sounds fabulous. We do that with clinical trials, et cetera, but not every variant across the space. We do reclassify though, basically every time a patient comes in we do a manual review even though we have a knowledgebase.

We pull that knowledgebase in and do a manual re-review of what’s there to make sure it’s current and accurate. If a variant classification does change and we have had that happen, then we will issue an amended report.

But what we’re trying to really encourage our clinicians to do is -- we’re trying to partner with them when patients come back and have progressed so that they’re reminded that there’s information in the record that might be applicable now.

It might not be -- if they're on therapy an amended report isn't going to matter to them if they're currently responding to whatever current therapy they're on. But when it matters is when the patient progresses and they're looking for the next option.

And so we're trying to figure out ways of identifying that from the medical records -- some key words, obviously, some change in imaging -- even an imaging appointment where you know that they're going to be restaged, trying to partner reannotation with those moments in time that might be most relevant to the patient and the clinician.

DR. PARK: I just wanted to add I agree with Karla and Kenna but I think ultimately what Kenna was saying about the somatic changes in the tumor -- I think that really relies upon academic medical centers and others with expertise because you need that clinical input to understand, is this appropriate for the patient to go on or off a therapy or start a clinical trial.

And ultimately I think if an academic

institution is going to commit to having a molecular tumor board, then they have to go in 100% and not just treat that as something that we're going to do so we say we have a molecular tumor board.

As I mentioned in one of my slides -- it's difficult, but you need a panel of expertise and you need people who are going to continually go back to the literature and really weigh the evidence.

And again, it's very dynamic -- it changes and it does make it difficult but I think, as I've said earlier, if you're going to do this you have to do it right.

DR. REHM: Yes there's another level of challenge -- that we've been sending out updated reports over the last 15 years and we've launched this system called a GeneInsight Clinic where when we approve a variant reclassification in any report that effects, an automatic email will be sent to the ordering provider who gets an update and a link to the new um, updated information.

The challenge is we think about this as a healthcare system -- we have providers who order tests based on a point in time that they're caring for the patient and they don't necessarily care for that patient for their lifetime.

And so just sending an update out to a physician who's now got this report feels some potential liability around what do I do with this information on this patient I cared for 5 years ago or a year ago, whatever, and putting them into a difficult situation.

So I think we really have to think about, you know, how do we support patients and their need -- particularly for germline variants where that variant may be relevant throughout their lifetime but their providers may change every month or every year or whatever?

How do we sustain a relationship where the patients are taking some active role in the ability to direct their own care, even if a physician changes -- and that's a really tricky dynamic that we all have to think about.

We also have to think about the fact that we largely don't get reimbursed for reinterpretation and how do we think about a reimbursement paradigm that supports the ongoing care and interpretation of information?

DR. MADISON: Excellent, thank you all for that great discussion. I want to open it up now for questions from the audience. I can always ask more questions but --

UNIDENTIFIED SPEAKER: This is not meant to be a controversial -- but I think it's an important thing to bring up. Um, Karla, you said something I think is very poignant.

You said if we hold on to pieces of information and don't update those on a regular basis we could potentially be putting patients at risk of harmful procedures or denying them care that could potentially help them.

And yet we're seeing it at the same time and I think all of us would agree to that statement -- that the quicker we can get information into the hands of clinicians who can use that information, the better we are to be able to advance care and that's what our patients want.

And when our patients sign consent to collect medical information they're expecting that information be made widely and be made in such a way that everyone can learn from that.

Yet at the same time we're seeing a very disturbing but a very real um, problem, where individuals are sequestering data -- they're not sharing data.

Now some institutions are publishing data and other institutions are calling those corporate secrets and are not allowing that data to come out. What is the role that we need as a medical community to decide what to do with sharing data as both private and public, you know, groups to help advance care for our patients? What's the role of sharing data?

DR. SHAW: My personal perspective on this -- so not speaking from AACR GENIE, is that I think that the patient's voice -- if that's truly what they believe that they're consenting to and I do believe often they are.

Their expectation is that we're doing whatever we have to do to make a difference for them and/or the future -- that's -- when I talk to patients that's what they say for the most part.

They have to -- if we're unable to convince our institutions to do so and they really feel that, then they need to also potentially decide with their feet where their care is and choose institutions that are proactively sharing and supporting data sharing efforts.

I think there's disconnect though -- between where a patient might go and their understanding of the level of data sharing that institution has from a research perspective.

I've never had a conversation with a patient where they've asked me, I'm going to go somewhere else unless you share my data. Okay, I've never seen a patient -- but I've never -- just to be very clear what my role is.

But we used to talk to a lot of patients in the consenting process, et cetera and no one has ever said that they would walk away from the number one cancer center in order to go somewhere else.

DR. BOWLES: So I think when we think about the concept of data sharing and uploading into public databases like ClinVar, I think we first have to understand that that is not a simple push of the button.

When we think about what information is available, how many variants we've seen over the last 25 years, we have probably over 60,000 variants.

And so to try and upload that into ClinVar or into any other public database isn't a simple 5 minute task. We're talking about thousands upon thousands of hours to get that uploaded.

And so we as a company have to ask ourselves what is the best use of our reclassification resources? And when we ask ourselves that question, our primary obligation is to the patients that we test.

When they came to us they expected a

definitive test result and we've made a lifetime commitment to those patients to do whatever it is we need to do to get their VUS reclassified so that they get a definitive test result.

And so we can either devote those thousands of hours into uploading data into ClinVar or we can upload those thousands -- or we can use those thousands of hours to develop novel, innovative, highly accurate reclassification tools.

And that's what we've done with things like Pheno analysis, mutation co-occurrence analysis, other automated haplo typing analysis and reclassification tools that we have developed over the last few years.

And you could see from the data that I presented from 2016 that resulted in 23,000 patients in one year alone receiving updated, more clinically actionable information.

And we anticipate thousands more patients to receive amended reports this year and in all of the following years. So understanding that reclassification resources are limited, we believe that this is the best way today that we can meet the obligation that we have -- that lifetime commitment to our patients to get their variants reclassified.

DR. REHM: I'm sorry, just to clarify as a regular ClinVar submitter, it does not take thousands of hours -- it does take work though. But I think we have to balance that work with the best interest of the patients knowing that, you know, and having done this now -- working with ClinVar, the value that I can add to my patients grabbing from the data, all of the data that's from all of the other clinical labs -- my patients are being treated much better with the massive data that we now have access to.

And I think that said, it is important to think about the commercial paradigm and insure that we have robust environments to sustain high quality services and just interpretation of variants is not the only piece of the high quality service.

You know commercial laboratories have lots of ways to provide and compete with themselves. A colleague of mine likened our environment to airlines and said, “Airlines don’t compete on safety, they compete on services.”

Do you get free luggage, do you board on time, you know all of the different things that are services the airlines provide, but we don’t want the airlines saying, “Well I crash less than you or I crash more than you.”

So in my mind in the laboratory interpretation business we need to share the evidence. The evidence is what allows us to provide the best care for patients and it’s not a lot of it out there for a lot of variants so we need to put it all together.

That said, the best reports, the best, you know, support for reimbursement and billing, the best turnaround time -- there’s so many different services that laboratories can compete with each other on and get better business and better revenue by competing on the things that are the service side of it.

However, we cannot deny our patients access to the actual evidence that will lead to the best outcomes in those patients in my opinion.

DR. BOWLES: And I think I need to go back maybe to a little reference that Dr. Park made about whether you teach someone to fish or whether you give them fish.

And I think where we think about sharing the data at Myriad is not necessarily just dumping data into a public database -- even with some of the evidence attached to it, but also what could we do to be advancing the science of variant classification?

So as we have developed these internal resources -- these internal analyses such as Pheno and mutation co-occurrence analysis, we have published that data, the methodologies. We have presented them at public meetings so that other laboratories, if they choose to, have the information that they can bring those techniques internal to their lab and have the opportunity to have their patients benefit from those technologies as well.

DR. PARK: So I'm just going to add a little bit -- that I actually believe everyone who consents to have their data in the public should be in the public.

That may be a little bit premature right now to think about but I do believe that we can learn a lot and save a lot more lives from data sharing and so I think if our patients are willing to consent then understand that it's going to be out there then that's the way that it should be.

The devil is in the details though -- how do you implement that and I think that's what Karla's getting at, to make it so that people don't get misinformed or misuse the data and do things that will actually be harmful to themselves rather than helpful.

And so I think those are kind of the steps in the roadmap that I see to actually be able to share big data and actually do more harm than good -- do more good than harm.

DR. MADISON: Alright, well excellent. Thank you all for this wonderful discussion and you all led directly into Session 4 which is going to occur after this 10 minute break.

We'll be talking about future directions for data sharing, standardization and establishing consistency, so thank you guys.

(BREAK)

DR. LITWACK: Alright we might as well get started. It's the last session of the day and so we've heard three great panels on the state of the science, you know, the cutting edge in interpretation and now we're going to look to the future in the last panel entitled, Future Directions for data sharing, standardization and establishing consistency in precision oncology.

And so we're going to kick this off with Dr. Dane Dickson who is the CEO and founder of CureOne and he's going to talk to us about that.

DR. DICKSON: Thank you FDA for inviting me to be here today. Um, when Dave said -- send me over a one line bio I said, you know, tell them that Dr. Dickson has a deep understanding and I was going to say usually what deep means when you come from the state of Idaho, is you're up to something neck deep that you don't want to be in.

So I have a deep understanding of payer policy and molecular genetics is what you can say. Today I'm going to talk about what are some of the obstacles for data sharing.

I think these are very real. I've been talking about CureOne -- what it is, what we're trying to do and what we see from a future perspective of happening.

So obstacle number 1 -- we cannot underestimate the problem that we face when it comes to sharing data that comes from the fact that currently molecular diagnostics are not reimbursed.

It's really difficult to share data if you're not getting paid for you know, even analyzing that data, then you're supposed to share it. And I recognize that it's costly to do and I also recognize that just because one individual payer may agree to something it doesn't mean another one will agree to it.

Next, um, data silos. This is something that I think we as a medical community really have to come to grips with -- that's why I asked the question in the last session.

The idea is -- is that I recognize that there are current business plans and business incentives and whole business models that are established around what data you have and I recognize that's getting worse not better -- especially as we get some big players that are common household names, multi-billion dollar companies are getting involved in this space.

But also academic centers, you know there's an idea that there's information that I need to have for publications, there's my intellectual property and sadly we're looking at revenue models -- how do we go through and sustain ourselves -- particularly in the molecular diagnostic arena where the equipment is very expensive and we've got to look at some way of trying to capitalize on that.

The third thing is let's say you do collect data but how do you know it's consistent? One of the things that I think is interesting is as we have looked at all these, you know, national or international databases, one of the fundamental questions that has not been brought up yet is how is the data if you were to compare it using the exact same data from one database to another database because you're looking at data that has come from an entirely different instrumentation, different methodology, different eras.

And so it's difficult to know, even without standards or without versioning, how do we know if the data that we're using to share is of the same value as it had before as it has right now?

And who's standards matter? Is CLIA enough? Well, in many cases sure. Is CAP certification when it comes to sequencing, is that enough? In many cases that may be enough.

Is it -- as the FDA has gone, does it require a New York State third party reviewer to get 510K clearances or FDA approval? We don't know what the right standard is.

And I think ultimately one of the questions is without clinical outcomes or clinical correlates, we don't know what standard is right.

We do know that a high analytic validity is necessary -- a high analytic validity is necessary if you are going to try to get high clinical utility -- we think that to be true, but we need the clinical correlations to really make this work.

And then I think one of the big things we have to look at is what does quality mean in this space?

Obstacle number 4 -- complexity and lack of evidence -- if anyone thinks that, I mean well, we give ourselves a great disservice in this arena of precision medicine by preaching way too early that precision medicine was going to solve all the ills of medicine and it was going to decrease cost curves and all these other things.

And we didn't tell people that the complexity of genomics is only the beginning. We've got proteins, we've got bio -- we've got the microenvironment, we've got all these other areas we are going to have to look at.

We keep on thinking maybe we can get data and we can put it all together -- this big data. And I hate the term big data because it usually means big mess or big pile of something.

And EHR data is dirty, incomplete, the data is not standard and then what's the transparency and sharing data -- how do we do that, that's an obstacle.

So enter CureOne formerly a group called the Molecular Evidence Development Consortium. We changed our name because Molecular Evidence Development Consortium was hard to say very quickly and changed it to CureOne.

CureOne is a 501C3 non-profit organization that was started to try to advance precision medicine by focusing on quality, evidence collection and transparency and putting it all together in a way that it could elevate the whole area of molecular medicine.

So we started with pulling together some of the top leaders in the nation. We put together people like Razelle Kurzrock and Keith Flaherty and Brian Druker and started talking about you know, precision medicine -- what does it mean?

We've got John Pfeifer who was on a panel earlier. We even have Neal Lindeman sitting down there that's now officially with us. We've got some good people involved with the project.

We put together -- started putting together data in the genomics committee of people saying, "Okay, how are we going to build something that could really advance the medicine?"

And what we did is we started saying on what do we need to do most and what we need to do is we really need to focus on the quality of what we're doing in the precision medicine and then the evidence associated with that quality.

Because you have to understand payer ease -- and payer ease right now is please, we'll pay you for quality -- not pay you for what you do, we'll pay you for showing that it's quality. So anything we talk about when it comes to precision medicine, we've got to say that we're looking for improving the quality of care of patients and we need to have evidence that needs to be transparent if we're going to speak in these terms.

There are other things -- accessibility coverage, standardization, shared data across -- all those things come into play. You know, patient protection -- it's a small little, you know area here, but I think it's important.

We need to remember that patients do want to be involved and they want to be informed when data is being shared. If you're going to start and you're going to say okay, what's a good standard I'm going to build something around -- there's the Agency for Healthcare Quality and Research, AHRQ.

It's a governmental agency we don't talk about very often but they go together and they say, look, if you're going to build things like clinical data registries, what should they look like?

This is their third edition -- they had to break it up into two volumes and the reason why is because they had so much stuff they wanted to say.

In chapter 22 of this wonderful manual and you should read it sometime if you've got insomnia. It's beautiful. They talk about something called the quality improvement registry.

And a quality improvement registry is this idea that you collect data not to answer a clinical question but you collect data so that you can then take that information to learn how to improve the collection of the data and what you're doing with the testing and you can build an iterative approach to how you're going to build the system.

In other words, it's a learning system. I won't call it a knowledge base but I'll call it a learning system that allows you to, as time goes on, improve what you're doing and how you're doing it.

So QIR -- a quality improvement registry has to have a few things that I think are important. One -- the quality, how do you determine quality?

I think most people will say that you know, if you're going to really show quality -- yes, you could can do good internal quality but if you really want to show quality it's good to have an outside review of some sort.

So CureOne said we need to have an independent group for example that would look at laboratory standards and we have a laboratory oversight committee that would look and take a laboratory's data, review it and say, "Yes, it looks like this laboratory is meeting a high standard."

We also put the other group of standards for clinical data elements. We held a meeting about 18 months ago with multi-stakeholders from the NIH and from industry and from ASCO and ACR and we had a few people from pharma and a few people from private payers and said -- what are the elements, what are the data elements you should collect -- and that was a paper that was published in cell just in the last two months.

We decided that you know, the highest value of understanding what the testing is doing is not tying it back to some existing database -- it is taking the information and really learning to say how does the testing when applied to a decision and a treatment and then an outcome, how does that end up taking -- what does that end up doing? How does that end up showing any benefit?

And transparency is the idea of saying let everyone have the ability to look at the data -- let everyone have the ability to review the data and publish off the data.

So we built a registry based on the HRQ guidelines and based on this multi-stakeholder group and it was launched official October of 2017. We enrolled our first laboratories in June, our first patient went in in October and we were enrolling and brought in about 15 patients and then there was a certain coverage decision that will not be brought up that decided we would slow down a little bit to figure out what's going to happen in that arena because we don't want to have to modify protocols several times over.

Um, what data did we collect? Well we decided that you know the complexity of genomics and the complexity of variant calling resides in multiple different layers. It resides in how -- what are you doing when it comes to collecting genomic information?

What are you doing when you are reporting the variants? How are you taking that information and reporting those to the clinician? So the idea was let's collect key elements of the genomic testing, and then let's collect high level treatment data on the patient -- what treatments were they given and did the patient respond?

And try to make them as simple as possible for the clinicians, try to make them as simple as possible for the laboratories but yet collect the information that would be necessary to really be able to drill down to determine why or why something didn't happen.

And the idea was to say let's put together -- and I won't spend a lot of time on this slide but the idea was put a bunch of laboratories together, put them in the same group -- have them go through the same level of standardization and then take that information and see how it's applied to one set of patients and see how those patients respond to therapy.

Then let's take that information, let's learn from it and then let's improve the standard of let's keep on moving on, keep on moving in a circle where we're continuing to improve the data.

We wouldn't allow any laboratory to come into this. We would say there needs to be a high quality standard that would be necessary for a laboratory to enter into the registry.

And our standard that we put in place was okay, academic centers, you're doing high level good quality work, how do you know that your neighbor down the street is also doing good, high quality work -- and it's very similar to like the New York State method of looking at laboratories.

Just one comment -- our registers we launched, one of the things we did late last year was this idea of saying look, if you're going to be a quality registry, you need to properly show that you can be a quality organization.

And so one of the things we did is we applied to Medicare through their innovation program to have the registry we're building also be able to report some quality metrics for Medicare.

And truly when Medicare, you know, has been saying we're going to pay for quality and they've told practicing clinicians, if you don't show that you're practicing some type of quality medicine you could lose up to 9% of your reimbursement in a few years.

So clinicans are panicking saying how do we show quality and what we said is well, if people are going to be participating in collecting data as part of a registry, we ought to go through and we ought to see if we can, by using that registry report quality metrics.

And lo and behold Medicare agreed to it. And so we had Medicare approve 11 -- these are only 6 of them and I won't spend a lot of detail with them but Medicare approved 11 of our metrics that if people participate with our registry we can help clinicians meet 11 of the 15 metrics they need to show quality and they can do it by participating with this prospective observational registry.

What do we see future directions of the CureOne registry? We hope that this CureOne registry becomes a great pre-competitive database of genomics that are reasonably transparent but have gone through at least some scrutiny of standardization, that have treatments and have outcomes based on those treatments.

And it goes into a database that is prospective and has patients that have consented to allowing the data to be shared. The reason why that becomes so powerful is that if a patient has agreed to participate, then it allows me to allow that data to go to someone else so that they can review that data or it also allows me the ability to drill down into an individual patient if I want to know more information or it allows us the ability to potentially, if there's any tissue left over, obtain a tissue specimen to reanalyze it using a different technique.

Well the idea is that this registry could act in this pre-competitive space to allow clinical trials to take place using other molecular methods. It could allow drug trials to take place because we have a national screening methodology.

It will allow academic centers to have a huge database from which they could you know, start their own clinical trials but also do publications and identify, unusual effects.

We also see it as ability to pull together other data sources, the ability to hook together let’s say with the payer database or pull together with a patient reported outcomes or for Heaven’s sakes, maybe at some point we call up Amazon and say we’d like to know the sales history or something else.

I don’t think we’ll get that far but the idea is the greater that we can bring in data that does not require physicians to enter that data, the greater we can potentially learn about patients.

But the problem is that we need to make sure that we are not -- that we are not, we need to not, not ask the patients to participate.

We need to ask them to participate and with that I'll end, thank you.

DR. LITWACK: Alright thank you very much and next we have Dr. Dennis Dean, who's the Scientific Site Advisor from Seven Bridges Genomics, thank you.

DR. DEAN: Good afternoon everyone, how are you feeling? Oh come on let's do it again, how are you feeling -- great. So I just want to first thank the organizers for inviting me here.

If you look at my title I started with creating a community view and that's because in my experience with working with the FDA and working on hard problems it's really about what's happening now -- coming together to talk about the problems and find the solutions.

Um, I want to give you a little background -- Seven Bridges is a relatively small company, a few hundred people in Kendall Square. We have a cloud platform for a distributing computation

It was mostly genomics to start but now we're analyzing lots of different types of data with a larger image database growing. And I think it's important to understand perspective so I want to tell you a little bit about my view of the world.

I started off in R&D and then I started managing our large projects, our Million Veteran Program, I worked on the blood pact and some of our FDA collaborations and most recently in the last few months the scientific staff in our Cambridge office reports to me and I've gotten an opportunity to look at what are the issues that are stopping us to doing great work.

And I started this talk by thinking about challenges and then I said -- let's change that a little bit. Let's talk about the opportunities. And so I want to talk about the places where I think we might want to invest some of our energy collaboratively and as a community.

Um, so what is the key problem here? So the key problem is, you know, and one day way back someone had some data, you had a programmer, they analyzed it -- it's much harder now because we need to have diverse datasets -- sometimes not located together.

We want to do analysis and database workflow management. We want to select a cohort. We want to do analyses and these datasets are so large we just can't do them on one machine.

So we need systems that allow us to do that. I have a computer science background so automatically I go we need a language or we needed a way of um, communicating the data.

And so one of the opportunities I had was to work on the BioCompute object that is sponsored in part by the FDA and the main idea here is that we want to be able to communicate NGS analysis.

And so in this room I don't have to tell you how complicated it is but I want to tell you why I believe we should invest in the BioCompute object.

The first is -- it's one of the first standards where there are descriptions of why you would want to use something, where you would want to use it -- including the error how did you validate?

And so the idea here is to create a standard where we could collect all the evidence in one place. And what I believe what's beautiful about it is that it's built in context and the context is how do we submit for approval at the FDA? And multiple stakeholders were involved.

And lastly, it can be used right now in a form,but it was designed so that we could use other standards under it so that we could expand out.

Now I want to talk a little bit about one of the standards so a common workflow language is one of the ways in which we could represent really complex workflows.

Um, and it was started in part with our work with the CGC so we could exchange between the cloud pilot -- but I'm going to argue a different reason why we should develop workflow languages and that is it gives us the basis for which to develop new tools.

And so this is outputs from one of our tools, Rabix, that takes a common workflow language and it allows you to view it from an application view to look at all the inputs and outputs.

You can look at it in a diagram or you can look at it in a code view. So the idea is we have to empower multiple people from multiple backgrounds and so that's one of the reasons I would argue that we want to invest more in languages.

I don't think I'm the only one that believes that so this is a blog post from Jeff -- I tend to use first names -- I hope that's not too informal. And he wrote this great blog post in saying how different languages come up to solve different problems but the long-term future will require us to resolve those language -- enable multiple languages to work together.

Um, one of the things that Seven Bridges was founded on was that we have to build tools that will scale and one of the concepts that we've thought about over the last couple of years is how do we build national scale analyses tools?

And one of those tools that we're working on and publication is just about to be released -- is a graph reference system and the idea here is to get away from the linear graph to a way of collecting all the variants in a way that it can be interpreted and analyzed and keep a population view.

I wish I could talk more about it but I can't so pull me aside I'll talk about it afterwards. Um, so I think -- and if you're in this community you saw when Deep Variant came out and it has huge implications for how we're going to think about data analysis.

And for those who are not familiar with Deep Variant I just want to give you the key ideas and maybe what we want to think about as a community. So it's really brilliant right? We're going to take the pile-ups that all our bioinformatics scientists take a look at and they call variants from looking, or we have tools -- and we're going to use that into basically really neat complex neural networks.

And they're going to call the variants for us. The bio-archive paper is really promising, but I would argue that although it's very promising we have to turn our clinical scientists and our computational hat on and go what does this mean?

So I think there's going to be a lot of discussion about this moving forward. Um, we often think about when we put these data systems together that we are going to do some analyses, but I want to remind us that when we put this data together are we going to be able to look at it differently?

So we have to plan for and think about how we bring in these new analyses. And this is something I saw at a U.K. Consulate presentation and I just found it really exciting.

So the main concept of this analysis was let's look at a polygenic genetic score -- so we're not going to look at one or two, we're just going to take them all and think about how we use them collectively.

And the presenter argued well, that it was a complimentary risk score. Um, so, I'm letting my epidemiology hat show -- although this is the idea of this conference is to think about making predictions for individuals, I also want to say that as we plan and build our systems we want to think about how we collect that data together to maintain an epidemiological view.

Um, so I am on LinkedIn every day and there is always a new ad for a new Apple app and I want to remind us that data is going to come to us differently potentially in the future.

Individuals are going to have access to their own data and their own ability to analyze and they're going to come better prepared than ever. And so we should think about that as we plan to move forward.

Just a few slides to wrap up -- I want to say we have more data than ever before. This is from Jerry Lee's posted presentation of two days ago. NCI has collected about 117,000 cases. How we analyze these large datasets will change how we do analyses.

So I've gotten an opportunity to work with Gil -- I mean I'm looking in the audience, a bunch of people have worked with him and he just designs really great projects.

And so here is his um, the Sync 4 Genes Project and the whole idea here is how do we bring standards together to empower groups to work together?

Once again I can't go through all the details but I think it's a great approach right? We develop a standard, we get pilot projects and we implement it and see what happens and we work on a very fast timeline.

And so this is my last slide but one close to my heart. When I was in epidemiology one of the things that jumped out at me was that if you looked at the race breakdown the outcomes were very different.

And so when we look at the data recorded for non-Europeans, we're doing better. But if you look at the percentages they're really slow, so this is African American, I think this is Hispanic here. So we just need to get a wider view of genomic information.

And I'm going to end where I started. I really believe this is about community. Um, last year I started with the BioCompute Project and it was just great to be with a group of people trying to solve a hard problem and so I believe that's what we have to do most and so thank you.

DR. LITWACK: Alright thank you very much and we next have Dr. Robert Grossman who is a Professor of Medicine and Computer Science at the University of Chicago and is the head of many, many projects and institutes which I will not go through.

DR. GROSSMAN: Thank you, so um, I built Data Commons, I built Open Source Data Commons. Data commons are used amongst other purposes to share large-scale cancer genomic datasets. I want to explain why their important -- the landscape for data sharing and what's changing.

I'm like a plumber. I build these systems for sharing cancer genomics data. Normally you don't really think about plumbers until your sink is backed up and dripping or your toilet is backed up so why should you listen to a plumber today?

I think the answer is pretty simple. With the scale in which we can work with data -- and this comes from the commercial cloud computing technology, we can do radically different things.

We can work with all of the cancer genomic data from agencies like NCI and other large scale collections of data and that allows us to benefit patients in a way we haven't' been able to do before.

So that's what I want to talk about. So I'm going to tell you at the beginning what a data commons is and um, that doesn't matter as much as I'm going to tell you why you should care what a data commons is.

The last several years I've been working with the NCI to build the NCI Genomic Data Commons. Today it has over 30,000 cases, it has cases from large scale projects like TCJ and Target. It has cases from Foundation Medicine. It will shortly have cases for GENIE and it's going to be part of an eco-system that's going to be able to uniformly look at the data from those over 100,000 cases that we just say from Jerry Lee's slide.

And I want to tell you a little bit about eco-system is going to be built. Um, the important thing about the genomic data commons is that it can work with large scale data.

It can work with the raw cancer genomics data and that's important because um, despite what the people who are building bioinformatics pipeline tell you, they work okay but not great and they're getting better every day.

So to have the ability to reanalyze the data when you need to with new pipelines when they're built is very, very important in terms of having the best data available to inform this sort of eco-system we're building.

If you're interested and you can't sleep at night go to gdc.cancer.gov put your favorite mutation in and see how the survival curves differ from all the patients in the genomic data commons for whatever clinical cohort you want based on the clinical co-variants for whatever mutations you want.

And it's really a good way to um, to get through the night. One of the important things that I mentioned is this notion of reanalysis. In general we've lived with the paradigm in which data by different groups, analyzed with different methods in different places, with methods that are usually not disclosed are sent together and then we get very surprised that when all that data is brought together it doesn't quite have the power that you might expect.

If, on the other hand, the data was -- the raw data was brought together, was analyzed with the consistent set of pipelines, was aligned with the same aligner, was called with the same set of callers and was recalled when you had better callers.

And that ability -- that ability to what is sometimes called “harmonize” the data is the sort of secret sauce of a data commons for cancer genomics data.

I want to tell you about another commons we built -- this is a private/public partnership that was alluded to. Seven Bridges is part of it as I think a lot of these companies are in the audience.

This is a data commons for liquid biopsies that was started as part of the cancer moonshot and importantly it’s able to sort of, interoperate with not only the GDC but other commons -- for example commons that are being built by NCI for Proteomics and images and this is done completely through a private public partnership with no government funding.

And what that means is that the partners can come together and work with FDA and others to um, sort of build a data commons for circulating tumor cells, cell free DNA, exomes, et cetera.

So what is a data commons? You don’t really have to remember this definition but it’s when we bring together data, large scale computing so we could reanalyze all the data -- and I haven't talked about it but the most important thing that's put over that are a common set of services so that you can identify the data with unique ID's so the data is citable, the data is findable, the metadata is discoverable -- common pipelines can be used and it can be reanalyzed.

And so what that change is -- you know the dirty secret if you're a plumber -- and if you're a plumber there are a lot of dirty secrets, but one of the dirty secrets is that most data is dumped.

People were embarrassed with that and wanted to give data a better name so they called that a data lake but most data is dumped and it can only be downloaded and it can't really be interoperable.

So what a data commons does is it starts with a data model and starts with these services like ID and metadata services so that you have a resource and that resource is open and can be used by other parts, by other systems automatically and people have built third-party libraries in Python and R and this is the beginning of the first step of moving from databases to resources that can be done at the scale for the data produced by large projects including Foundation Medicine, GENIE, TCGA, Target, et cetera.

I want to talk a little bit about the -- oops, is my timer -- I forgot to set my timer. You're going to tell me when I'm out of time right? Okay, I want to tell you about the data sharing landscape in precision oncology.

Um, I haven't talked about this but there are two things going on. First of all data sharing can be done safe and compliantly. It can be done where if you create a resource and not a dumping ground or a silo, but if you create a resource the data could be left locally and you can still have some interoperability.

So on one side we have the ability to protect patient information. Importantly, on the other side, we have the ability of patients to benefit when large amounts of information are brought together so that we can get new discoveries.

And so that balance is what's at the center of a data commons and we heard in just a previous talk about - the deep learning techniques that are beginning to be applied.

They work -- well in PowerPoint they work whenever someone wants to give a talk. But in real life deep learning works with a lot of data and we can't have a lot of data unless we build an eco-system like this.

The data sharing landscape is complicated. You have got these swim lanes, you have the basic discovery, you have got the clinical trials, you have got patient care, you have got quality and safety, and we need to interoperate these.

And you know, I think one of the things that I hope emerges is when you have the strength of evidence databases and they can transparently reach back to systems like the GDC and BloodPAC and GENIE with enough identifying information that protects patient privacy with the techniques we have now so that we have the patient level and case level information so we could get the enrichment that we need.

One of the exciting things that I think happened is in the end, if you're a plumber you like to think that what you do when you build systems is what's important.

It really isn't. There are lots of plumbers and all they do is going to build you a system. If you get the right plumber the system will work. Most systems don't work because they're not that many competent plumbers but that's a different story.

What it comes down to in data sharing is what are the incentives -- and there are very, very few incentives that work and I want to talk about some of the incentives.

One of the exciting things is the International Committee of Medical Journal editors got together and put incentives out there for people when they publish about clinical trials that a certain amount of information be available.

I'm going to come back to incentives but I want to sort of, talk about something that's really important. Um, we heard about the papers that a lot of that information may be outdated.

Sometimes that information is not reproducible and what's changing now if we look at how we go from target populations from study cohorts to samples and we look at the replicable research from single lab to multiple labs -- in terms of going from samples to data -- a data commons allows you to go from data to results in a reproducible way because we have techniques like common workflow.

Importantly, we have the ability to re-executive the pipelines, to share the pipelines in whatever way and so we can uniformly process and harmonize data.

We can run new algorithms and in this space in which almost all evidence is weak, you can accumulate evidence so the unit of progress is not a paper that may or may not be right -- the unit of progress is the accumulation of data in which could be reanalyzed.

And so we can constantly reanalyze the data so that the weak effects that are so important in cancer can emerge whenever we do a reanalysis.

So I want to come back to the incentives and the guidelines. So, you know, one of the most important incentives if you're publicly -- if you're funded by federal agencies you have to share the data.

Now, what they don't talk about yet is we haven't still quite figured out how we fund the infrastructure and have a sustainable infrastructure and some of the earlier talks talked about that -- how do we get a sustainable infrastructure for sharing, but at least we've taken the first step.

Um, where we are just beginning to make progress is the private foundation which funds about 25% of cancer research each year are also beginning to require data sharing and they're also beginning to think about how they could provide the infrastructure so that could be shared.

We talked a little bit about the strength of evidence and so on. I think the other lever we have is the payers, in terms of providing the best us of their dollars and the maximum return of patient benefit, can begin to require sharing so that they could be also part of the eco-system.

And so I'm thinking of how do we make this transition? And so to my principle -- which was kind of similar to the panel out there, is as an organizing principle, we have a data sharing model that requires free access to data but then encourages the competition and sustainability around all of the things you do around data.

How do you build software? How do you put the professional services? How do you do the value of added products? How do you integrate with the EMR? But the basic idea that the raw data and the processed data -- you should not be able to be reimbursed unless you share that basic data.

And so, common support -- lots of data sharing models -- I'm not going to go into it.

Um, so I think we've seen something major with the emergence of a data commons.

We have the ability to reanalyze data, to do it at scale, to have common data models, to have harmonized data and to be part of an eco-system and if we tried to build one big system, it's not going to work.

If we try to build a reasonable eco-system it can. And so this is my last slide. I think there's something new that's emerged in the last several years with data commons.

I think they're a platform for open data and reproducible data that will benefit patients andI think -- a lot of the work I do as described is supported by 501C3 not for profit corporation called the Open Commons Consortium whose job it is to make it easy for you to build data commons so you never have to worry about plumbing, thank you.

DR. LITWACK: Alright well thank you very much and I'd like all the panelists to come up now. So thanks very much for some great talks and I thought we would start off at the other end with

Dr. Lichtenfeld who's the Deputy Chief Medical Officer at ACS and I thought he could introduce himself and say a few words.

DR. LICHTENFELD: Thanks David and it's really a pleasure to be with you and it's certainly a pleasure to be able to speak to the folks who have been here for the entire day and I thank your attendance and hopefully those that are watching through the webcast is a testimony to what really has been a fascinating discussion of a very complicated topic.

I move into this discussion with some degree of trepidation and the reason for that is I know what David wants me to talk about and then I always wonder what am I actually going to talk about and it may be a little bit different.

I certainly have experts who have preceeded me who are much more familiar with some of the more complex data requirements in this discussion we've had today and I can't help but reflect on the fact that I sit here not only as a member of and on behalf of the American Cancer Society but also as a patient advocate, and frankly as a patient myself.

And I can't help but wonder about some of the things that we've heard today and if you want to get a capsule view go check on Twitter under the hashtag FDA Cancer Variants, I think -- is that right Hisani, you know what I'm talking about.

My thoughts are there and you can see them. So I'm going to talk, perhaps off of that. We really, you know, I'm going to say this again as I mentioned with some trepidation.

We are in -- as you all know, you're professionals in this field for the most part I'm sure -- we're in a rapidly changing environment that is impacting the lives of those we serve and from past experience sitting here today to serve to reinforce--we need to find a way to make this work for the benefit of those we care for, those who are currently our patients and those who are consumers and will be our patients.

And we also have to bear in mind that we have to be able to offer those who care for our patients -- the clinicians and all the folks who are involved in that process, we need to serve them well also -- in addition.

I'm reminded of several things so we need a more flexible legislative, regulatory and payer process than what we have.

Some of the things are so simple that we haven't even talked about or thought about -- I don't know if you're aware of this but under Medicare rules for the most part, a woman who happens to have evaded testing for germline mutation who presents to her knowledgeable clinician at the age of 65 who takes a family history which frequently hasn't been done with any degree of accuracy prior to that time and has an absolutely convincing history cannot get -- cannot get screened for BRCA because she doesn't yet have a disease -- gets breast and ovarian cancer, she's in the door.

That's not just for women but for men as well. So when you start from that position you begin to understand the difficulty that one has. And then you translate that to what I would consider and I don't mean this is not pejorative when I say -- this is just a fact of reality.

When you sit and listen to everything that's been discussed today which is incredible science, very high level, by people who are at the top of this profession, at the top of the science, the top of this technology and then you try to think how do I translate that information into the care of a patient in the city that I used to live in in south Georgia town -- that I lived in in South Georgia?

You know that they're pretty good. A whole lot of care in this country is administered outside of major medical centers and when you talk about molecular tumor boards and when you talk about some companies -- not all, we've been fortunate to have high-quality companies talk to us today and high-quality university labs, but not everybody fits into that category.

And not everybody makes that investment in terms of um, oversight, updating, notifying --

these are really significant issues. How do you build an eco-system of trust? How do you build an eco-system that provides the information that people need?

How do you build an eco-system that promises to people that the latest information -- the latest information will be available? At the moment it's available so that we don't have a situation for example, as came up a short time ago, of not updating databases so that somebody may get a germline test that has a newly proven variant to be of significance and not find out about that because proprietary information or because we don't have systems in place to update for six months, a year or whatever.

So we have a lot of work to do. We need to have some assuredness, we need to have certainty -- we'll never have certainty, I understand that, but we need a better system.

We need to start thinking about the people we serve -- the patients we serve. Sometimes I think like with the ERH and you're talking about massive databases and I think about all the EMR situations that I have -- another area that I get to talk about like this.

We really need to be able to think about who's at the center of this discussion? Who needs the information? How do we get it there? How do we have payers that are responsive? Is it going to be through coverage with evidence type of processes and not just with CMS but who's going to pay for this?

This is a living test. Who is going to pay for that infrastructure that was previously described and I think is terrific of making sure that even though that test was done previously we're going to update that information.

We haven't thought about that either. So we have a long way to go. I, as a patient -- I as a consumer, want to have some degree of certainty.

I as a physician, as a clinician -- want to have some degree of certainty that when I make a comment to the people I serve that I care for that it's accurate, that it's meaningful, whether it's actionable we know what to do and if it's not actionable, we know that as well.

And in closing I will say that following our discussions today, you know, I have certainty that if I go to certain places -- some places, I'm going to get that.

And I also have -- unfortunately, some certainty that if I go to other places I may not. So with that David I'll turn it back to you and hopefully we'll engender some thoughts about the topics.

DR. LITWACK: Yeah thanks, those were great comments and you know, so one of the nice things about this panel is I think we're really spanning the whole eco-system here from merely basic informatics all the way through to the patient and I think that's, you know, the sort of communication that we need more of.

I did want to start by just asking, you know, because you all come from a somewhat different perspectives -- if you had, you know, sort of along the lines of what the barriers you're facing

-- we've heard a lot about barriers and there are so many things we could fix or work on from the standpoint of where you sit, what would be the most important thing to address other than payment?

I think payment is obviously, you know, the easy answer for that.

DR. DICKSON: You've got to look at the whole eco-system and say what is going to allow individuals to want to share data?

Um, you know, because it doesn't matter how nice your IT tools are, it matters you have got to have to get the data that comes into it. And so if we say okay, what are the incentives for physicians to give up data for example, what are the incentives for laboratories to give up data.

I think we've got to start saying, "Okay, David I'm sorry but you can't take payment out of the equation if you're looking for laboratories who want to share data," -- I think the CMS decision that's pending right now with the coverage limits development is absolutely one of the best things that's ever happened to advanced precision medicine, depending on what the final comments are.

But the physicians -- it doesn't mean the physicians are going to collect data. So you have to say how -- what can I give to the physicians to allow them to collect data?

Well one of the reasons we started CureOne was so we could do a couple of things. One -- we could help them meet the criteria for quality that they need to meet already and so Plus 2, as a non-profit organization, we can potentially incentivize them to collect data through some appropriate reimbursement that's less than fair market value which we could do which a laboratory couldn't do by themselves.

What else could we do? We could incentivize academic centers by saying now you've got access to a large database that you can publish upon or pharmaceutical companies -- now you have a large access to a database that you could go through and look for variants or other things.

Identify what those incentives are and helping those incentives to take place. And it's not so much that we need to create an eco-system, we just have got to put all the pieces together.

DR. DEAN: Two areas I wanted to address -- so if I could fix anything it would be being able to associate context with data. So if we want to share data -- we want to be able to share it throughout the system and so right from collection all the way to clinical analyses -- and so that's really about context.

And the reason why I say that is because there is value there, right? That's something people would invest in. I trust the data as you said. It has the information I need.

And I would also say -- because it's expensive to do right -- because there's the big infrastructure?

I would argue in my experience in working with the FDA is that they're really great models for doing that and one of the ones that comes to mind is my colleague, Ogan, in the back is working on this C2/C2 project and Wemming runs the project and he got companies together to invest -- it must be a million dollars in sequencing and analysis because it's the right project and everyone wins if we work together to do that.

So I think two things. One is we have to be better at sharing data that has value and two -- we have to find creative ways to fund.

DR. GROSSMAN: I think it in the end it's going to come back that we have to rethink this balance between protecting the patient and the rights that patients have to good quality healthcare.

If what we are looking at in precision oncology were easy -- if the effects were large, then we wouldn't have -- we wouldn't need the data, we wouldn't need the data sharing.

Now, as we improve our understanding we may get um, you know, the ability to do inclusions and exclusions for certain homogeneous questions the effects are larger than they are now, but we are going to need to do this at scale.

The incentives we have -- it's really if you share data, there's a lot of liability so I think we're going to have to come back to addressing the liability.

We don't talk a lot about it enough but the only thing we know about large collections of data is they eventually get breached. So if we look at this that with the current set of policies we're going to lose -- we know where we have to go to is at scale we can understand weaker effects and benefit the patients.

So I think the equation has to change and a lot of people have talked about this or alluded to it today but in the end the patients are going to have to drive the sharing and we're going to have to go to these eco-systems built by patient partnered research where there -- in a regulatory environment, provide the ability for the data sharing to be done with the same tools we have today but at the scale we need so that they can get the benefits they can.

And so I don't think those are major changes but we're going to have to be creative at the legislative and policy level to get there. We have the technology. The patients have the desire -- especially the sicker they are, but that leap to the data share -- and I think that's the next major leap.

I'm pretty sure it's going to happen in the next five years but I think that's going to be an absolutely critical ability to improve patient outcomes.

DR. LICHTENFELD: Well I'm going to just move on to what you all said because I'm sitting here thinking and clearly data acquisition, data analytics is important. I think and I do believe that many folks, if given the opportunity -- and this is not a new thought.

Raise your hand would you be willing to share I think would be willing to share. They want assurances regarding privacy. I have seen some horrendous HIPAA, I have spoken and written about it,I've seen some horrendous HIPAA permissions that allow them to take whatever data they have and -- someone has, an institution or entity has a healthcare system to share with whomever, with no recourse.

So I think they need assurances plus we unfortunately -- and one of the points that was made earlier I think is worth reinforcing, there are communities within our nation that have not been treated well in the research enterprise in the past.

Those situations live on and they influence the willingness of some communities to participate. If we're going to be effective we need to have this as broad an effort as possible. We need to reach out to those communities and engage them.

We need to make sure they are included -- that there's every opportunity to be included and we have to take every precaution and protection possible that everyone is treated fairly.

What I don't think -- what I don't think is appropriate in this rather large enterprise is that we run into a situation again where somebody's individual data becomes um, so unique that it becomes someone else's business case.

Uh, we've seen that happen in the past. Books have been written about it and I think that we need to get a mindset that the data really needs to be in the public domain and I think we need standards so that that data -- I've talked to Dr. Gross, but you probably don't remember -- I asked him a question one time at a meeting.

I said with all the difficulty in standardization of information about patients, how do you get around that? And we certainly know that some of our sister organizations have had difficulty doing that with all their good intentions to get patient data, so we need to make sure that those standards are in place that the data in fact -- not just the genomic data but also the clinical data can be queried and useful in evaluating what we find from the genomic information.

DR. LITWACK: Alright, thank you. So obviously a lot of your comments were around data sharing which was sort of the next set of questions I had teed up and you've answered many of them.

But let me just ask you from your experience what you think would incentive data sharing? We've heard some ideas today but I would like to get the opinions of this panel, go ahead.

DR. LICHTENFELD: Can I throw something out there that I wrote down on my notes a little while ago and something that doesn't get thought about a lot and maybe you are familiar with this.

But I actually found the CureOne comments instructive when you talked about going in the MIPS program and using that as an incentive.

There are organizations like the National Quality Forum that have quality requirements and they did move the needle with some of the more now considered standard tests but it took some effort to get there and even things as simple as estrogen receptor and HER-2 and things like that -- now we're almost at 100% and that's really quite an accomplishment and that was done through quality metrics.

So in the sense that there's evidence base for doing some of this and making sure that it's implemented in a quality metric that can be measured, that becomes incentive.

Not everything we do -- and money is important and follow the money, yeah we all know that, it's the Sutton's law but the reality is there are other metrics and quality metrics if they can be embodied into this, that is in fact it is a quality -- an indicator of quality that we are engaged in this process and that's recognized as such -- that would go a long way towards getting the attention of clinicians to participate in these types of efforts.

DR. GROSSMAN: I think it's pretty simple. If people have the option to share data no one shares data. So unless it's a requirement you don't share data and I think the flip side -- just to build on what you said is in several years, you know, the patients have to control it -- including the ability on a go forward basis if they change their mind about certain data sharing, they have more knobs so that they can stop the sharing of a certain type and change the type of sharing.

And we can give those knobs to the patients so those are the two sides to it. In the requirements, you know, we've talked about, you know, the people with the money in the end are going to make the requirements.

DR. DEAN: So I thought I'd start with what I thought was an exciting story. We have updates every Monday and one of the people on my team came in and said, "We had a grad student analyze 20,000 samples over the weekend."

And I was like whoa -- really? And the answer was yes. And that's because of the investment of the NCI in the cloud pilot. So we have three cloud pilots and what's interesting to me about the cloud pilot is that you have to think differently on how you engage them.

You have to upload your data somewhere and then you have to trust that it's going to be computed on the cloud. And so part of where I'm going with this is part of it is that we need a new generation of trainees who are comfortable with working with the data in a different way.

We have to build the infrastructure -- so that's one aspect. I think training is going to be important.

And then the second part, Vahon isn't in the audience so I just want to briefly mention one of the ideas he's been talking about and he's in the FDA.

On a health exchange network where we incentive making data available and having an infrastructure for paying for those additional services and it's going to require an additional investment but I think we worked the whole system.

We work with the grad student and we work with the large agencies to build the system.

DR. DICKSON: One of the things I think we forget is to quote one of my friends who is -- was a CTO for many years -- he said it really doesn't matter what software you have if you can't aggregate the data in the first place.

And so some of the things that Bob is talking about is so crucial about and you know, Dennis are talking about are so crucial in getting the data there.

But at the end of the day it's physicians that have to give the data because unless you can, you know, spend the money to access the EHR you can then do something to clean the data in the EHR and find the information in the EHR you really want and I won't even talk about whether or not the patient has a consent or has signed a consent to let the data be shared.

I mean I won't even go there. But if you don't understand what the physicians need in their workflow and the problems that they're going to face in sharing data -- just to go into a physician saying, "You should share data."

You know, I see patients two days a week in rural Idaho. I know what it's like to be given your latest mandate because someone says here's what you need to do you know, thou shalt report, thou shalt do this, thou shalt have this program, thou shalt fill out paperwork in this way.

And if you don't we're going to audit you and we're going to pull -- we're going to claw money back. And so I mean, that's the environment the physicians are facing every day.

You also have patients that don't know what's happening with their data and patients are sitting there -- we are in the post Henrietta Lack's era. We can't be so naïve to say um, we should have more liberal data sharing when most IRB's would say, "Look, if you're taking data and you're sharing it for commercial purposes, the patient should have been able to have the ability to agree to that or not."

And so I think we've got to just understand, you know, what is happening with the patients? What is happening with the physicians and really focus on what can we do to incentivize them that isn't another stick?

You know, what could we do to help them be able to improve the care? What's their incentive to be able to do that?

DR. LITWACK: Okay so I want to go -- just to build off what you said Dane, you know one question I've always had is -- let's say you have the informatics infrastructure to share -- for every patient in the U.S. to share their data.

And I think about, you know it was mentioned -- one thing about big datasets and this is well-known is that they get breached. And I think the question is how much trust do you feel like patients have right now in the system and if we could allow patients to consent to share their data -- every patient -- how many, you know, how many do you think would?

DR. DICKSON: I mean I can speak from an oncologist who sees patients. I mean most of the patients -- at least in Idaho are pretty conservative proper type state.

They, you know they are very happy to say if my data can be used to advance care for someone else I'm willing to do. Um, you know, when we consent a patient we have another consent. And, by the way, we are putting together appropriate, you know, safeguards but we could have a breach.

We're telling the patients that upfront because we can't guarantee that someone can't breach it -- we'll try to do everything we possibly can and so I think that ultimately transparency becomes a big issue with patients.

And if they have that transparency and the hope that their data really is going to go and make a difference, you'll get probably 90 plus percent of the patients that are willing to share their data.

DR. DEAN: So I would agree with that. One of the first things I did when I started working on the Million Veteran Program was read everything I could find.

And I was really surprised to find back in 1998-2000, the VA convened workshops with the veterans. And by 2000's the veterans overwhelmingly said we want to help our veterans with our genomic data.

And these are military professionals and you may think they're educated or not educated but they wanted to help their cohort. And I think that that's the case for most patients.

Um, the thing that I find interesting -- and I want to make this two points. I like stories. The first thing is I think it's about policy. We have to think about what are the policies that incentivize the sharing so that happens.

Um, and I'll leave it there -- I have another story but I'll save it for after.

DR. GROSSMAN: I changed -- I mean I think sharing happens at multiple scales. The best sharing is transparent because it's a side effect of something that's been built that directly helps the patient.

It's going to take us a while to get there, especially for EHR data. Um, the way I think of where we're going with sharing is for many people building on what Dane just -- I mean what Dennis just said -- it's about sharing with people they care about.

They care about who they identify with. They typically identify with others who have the same diseases they have. And so if you're sharing with people who have your same disease in a patient advocacy group, you're willing to do much more sharing than if you're just sharing abstractly with the general patient population because you don't identify with that.

And so, you know, if you look at this from the viewpoint of the patient advocates, they look at -- they want to benefit the other people who they identify with in those advocacy groups and they're willing to do that and I think that's how we have to think of it -- not in some abstraction.

We may build systems in the abstract, but the sharing has to help with the people that they talk to every day and identify with every day when they're trying to figure out how they're getting through this.

DR. LICHTENFELD: I agree with you absolutely and there are certainly advocacy organizations out there which have done that and there are advocacy organizations out there which would like to do it.

And I do think it's in the public good and the public interest to advance that. I think the other party that's not been brought to this table and discussion are the private payers.

And um, I don't know if there are any payers. I'm not asking you to identify yourself in the audience or watching but they could play a much bigger role. They have a huge problem.

They're trying to figure out how to pay for it. They're trying to figure out what to pay for and they -- as a general principle although they're asked to pay for participation in clinical trials or coverage in clinical trials -- let's just say they're not willing participants in running to the table to do that.

And they're also not necessarily -- and this is not -- I'm not being critical. I absolutely have a point of view which you can tell but I do think that just like CMS has coverage with evidence opportunities I do think the private payers can do better in that regard too.

The numbers that I heard today and I may be wrong so correct me -- the number that I heard today was that 32% of the patients, or -- about 30% of the patients who were tested have an actionable mutation.

10 to 20% of those folks -- actually 10 to 25% I think the numbers have varied a little bit, actually got benefit from having had their genome sequenced.

And those are pretty substantial numbers folks, from my perspective. And if you want to get attention quickly then you say, "Okay, if we're going to do this test and we're going to get that information, we do want it shared -- not unlike TAPUR or MATCH, whatever.

But we're going to do it at scale. We're going to make your lives easier. We're going to have the FDA -- although it's made great efforts to try to accelerate the you know, single use IND's or whatever, they're going to try to make that as easy as we can and you will benefit if we find something important.

I would venture to say for somebody who has an advanced cancer -- a 1 in 3 chance of having an actionable result from doing that test is a pretty high incentive.

But as mentioned previously, when you're the clinician taking care of that patient and you're faced with a lot of bureaucratic issues and then you don't know if you're going to get paid or not because the insurer may or may not decide it's going to be a payment issue -- that is a big disincentive.

Let's work on making the incentives incentives, and let's move forward with collecting the information. We need to do that and it's going to be much more robust going forward than it is even today, thank you.

DR. LITWACK: Alright well thank you very much.

DR. LICHTENFELD: So another part of this discussion is not just on data sharing but on standardization and consistency.

And we've heard a lot today about sort of the cutting edge of somatic variant interpretation and practices and the different things people do.

And so I think one of the questions I had was -- what's the impact of that variability? How much do we need to get people to standardize what they're doing for all this -- for this effort?

And how much do we not want to constrain what people are doing at the same time? I mean, is there a balance between standardization and sort of the ability to develop?

And so I was going to throw that out there for whoever wants to answer.

DR. DICKSON: I mean the Holy Grail right now is clinically annotated genomic datasets. And Bob, in the genomic data commons how many records are probably clinically annotated? And so we've got some records that are getting some clinical annotation but the question becomes is where are we going to get the rest of the clinical annotation?

Is the clinical annotation standardized? Can we go through and get that? So without, you know, when you're getting it from a clinical trial where you've got the standard, you know, methods of collecting data -- it's much easier when you're looking at a patient that's living in a community center somewhere.

And so I think that when it comes to clinical data standards we've got to look and say okay, what are the most crucial data standards?

Things like, you know, treatment? Things like response? Things like time on treatment? We may not say that you know, a non-drug altering toxicity may not matters -- sorry FDA, but maybe a grade 2 or even a grade 3 in toxicity isn't so important to understand if it didn't change the therapeutic decision-making.

And so, I mean we've got to be able to standardize that data. It's not in the EHR's and so we have to decide -- either we are going to, you know, make physicians -- and I use the term "make" purposefully.

We're going to make the EHR such that a physician actually has to say the patient has had this type of response of this portion along the way or we have to have the physicians fill out forms, you know EHR's -- excuse me, case report forms, CRF's that do have standardization in them.

It's the clinical standardization that is so difficult outside of clinical trials.

DR. LITWACK: Um, the older I get the simpler I try to make things. Whenever I get a complex question like that that I don't understand I think about banks and bank tellers.

So if you think about how a bank works, you know, for a long time banks complained that you know, it took too long to get money out. And so they didn't sort of set standards around, you know, how the bank tellers pulled the money out of the drawer and dealt the money, they eventually came up with a system that they had an ATM that did a certain amount of things completely automatically with the standard leaving the bank tellers to sort of do the rest.

And I think that's sort of where we are here and it's building what Dane and Dennis were saying is, you know, we can transparently put standards around a certain amount of the back end of this that will make it easier so that more time could be put around the things that are actually quite hard right now which is the interpretation in and around that.

So when we have something like this I really don't think we should try to fix the whole system. I think we should look for the ATM portion of it that can make the rest of it quite a bit simpler and put a lot of standardization around that.

DR. DEAN: If I can build on your comment. I wrote down two words. One is that standards have to have added value. Sometimes we think just about the communication but we want to provide something more that enhances the medical experience.

And the second thing is that standards and abstract aren't very useful. We need something that's very usable and easy and so transparent that you didn't even realize the standard was there, so we have to really invest. It will take some investment to make that happen.

DR. LICHTENFELD: I'm sitting here trying to think about which standards we're talking about okay?

DR. LITWACK: I forgot -- that's another question too.

DR. LICHTENFELD: Fortunately not for EMR's and you know, but for my simple brain, that's the standard I'm thinking about. I'm thinking in a little bit different context. I will tell you that coming out of this discussion today I am very concerned about how much the typical clinician caring for a patient really understands about these tests.

And going back to the analogy -- the analogy that I use is a number of years ago um, my son tried to take me -- you know, we were in Baltimore at the time, tried to take me to a school class to learn how to program an Apple Lisa or Apple 2 -- one of those things.

And I said, "Kiddo, it ain't gonna happen." And when I finally got the Packard Bell that came out of the box, I could press the button and do word processing I was a really happy guy because it had utility to me.

I think, you know, the medical community needs to have an assurance and an understanding that the information that is being provided is in fact accurate, that it's up to date and that it's actionable and it makes some sense, okay?

We're not there yet and I'm going to leave it to all these folks to get us there -- all the people who spoke here today. But we have to have a certainty that it means what it says.

I have to share that I had the opportunity -- maybe about two years ago -- I haven't looked at it more recently, to look up a guideline on prenatal testing -- anti-natal testing for genetic variants.

And I will tell you I have a wife who is not listening. She was obese, she's now just GYN and I went to her and said how much of the specificity and sensitive do you know about pre-natal testing?

She looked at me cross-eyed. She says I'm drawing the blood test, they tell me it's right or wrong. Well if you read the guideline that came with the test -- it was 30 pages long at that time so this is a society of reproductive -- internal female medicine had a guideline.

It was full of information about the limitations of the test. You know what it meant, what it didn't make -- and placenta is a pretty big organ right? And here we are looking at, you know, now cell free DNA is little bits and pieces and you're trying to figure out sensitivity and specificity not withstanding with the excellent report last week from Hopkins.

Clinicians need to know what it means. They have to have confidence that what they're being told is accurate, they have to have confidence that it is actionable -- those are the kinds of standards I'm looking for.

And so I want data. I want analytics, I want to know if something new is found and if it really -- in particularly, if it -- I don't care about some of the side stuff, particularly if it makes a different in patient care -- those are the standards I care about, you guys figure out how to get us there.

DR. LITWACK: Alright so let me say just one last question, but I like the ATM analogy because I could also solve the payment problem here. Um, but um, I just want to ask -- so we've talked about automating the things that can be automated.

And so that sort of leads to my last question which is what can be automated here? In particularly I was thinking during the talks today we're still talking about really mostly expert humans sitting down and curating evidence -- reading papers, meeting, being on phone calls, there are so many variants out there.

There's sort of, you know, this is where you might view that -- so this is where there may be a real bottleneck. So the question is what role do you see automation -- particularly AI, machine learning -- what role can it play in the future here realistically?

You know I think it's still been a little iffy you know, what it can contribute here so.

DR. DICKSON: So you can't analyze data until you have data. Um, the healthcare industry speaking from, I won't name the EHR battles that are going on out there, but there are EHR battles whereas there is a standard ATM, you know, protocol.

I can use my card anywhere across the world or my chip just works, even though someone may steal that chip Bob, but I can still use it.

The problem is -- is that EHR's don't communicate with each other and then neither will they ever. Academic centers don't communicate with each other, neither will they ever unless things change.

In other words we cannot decide on one standard like the banking industry did or something else. And so one fundamental thing we need to do is be able to decide how are we going to communicate with each other and until that I don't see any widespread use of AI or widespread use of implementing some other type of, you know, wearable technology because once again it's going to be proprietary -- it's going to be in a small subset.

Someone is going to ask -- what's in it for me? I mean how am I going to, you know, make money off of this and that's where it becomes a problem in medicine is everything is so fractured because of concern for anyone of a number of incentives for companies.

DR. DEAN: So I think I mentioned to people beforehand I did my -- I was at the Brigham for a number of years, did my Master's in machine learning with a bunch of physicians. And when I could finish my Master's one of the things that came to mind was that what we really need are tools that assist the expert so I would be very wary of any system that said we can replace the physician.

And that's because the physician is trained to um, to redirect when they get a little piece of information -- one word from a patient's parent. So where that information is going to come from isn't clear and so we can't build tools to take in all those inputs.

So I would first say let's work with AI for the assist. The second thing is we will learn by doing that -- what we can automate and what we can't. And what we can, we do. And then for the really hard problems where we don't have the data for -- we just can't do it and we just have to acknowledge that and find a way to get there.

DR. GROSSMAN: Before I was a plumber I used to build machine learning models. I look at this pretty simply -- we live in one of two types of worlds.

One of the lessons we had with machine learning and what we call it now, AI. But one of the lessons we had is instead of building more and more complicated complex models, um, on small data -- you're almost always better off building simpler models on larger scale data.

And so right now instead of building more and more complicated machine learning models over all these papers and studies on 30 and 20 people that were done with lots of motivations to get them published -- either at scale when we have the core diagnosis, no matter how bad it is and how nosy it is and the sequence level -- the exome or targeted panels or whatever we have -- either something emerges at scale in which case we can understand what's going on for that patient or it doesn't.

And if it doesn't -- I mean nature has always been more complicated. You know we created the whole notion of junk DNA just to explain the fact that nature was doing something with that.

So either we have the data at scale we need so we can make simple observations -- and I think we will and we'll get certainly more, or nature has been clever again and we're going to need a whole other scale of data in which case we should just wait and not try to go back to the complicated stuff on poor data at small scale.

DR. LICHENFELD: I don't want to get too far off topic but Dane said something that I'm actually -- I'm going to, I have to put in my two cents on this one.

He said the EMR's -- the EMR will always be the EMR. And I will say that in the world of disruption and you start thinking about what's called the longitudinal record and block chain potential -- I can't help but think of one word like Amazon, you know.

So I'm not convinced that it's always going to be as bad as it is right now because right now it is bad. So having said that I'm going to take a little bit of the contrarian view -- and my contrarian view is that um, again let these experts talk about automation.

I do feel there's going to need to be a learned intermediary which I think is the term of art in the FDA world -- a learned intermediary to help us interpret and make these tests understandable in terms of the application to patients.

I don't want to have it go unnoticed that Dr. McLeod made a comment about a personal -- about the genetic counseling service they actually have -- you may be aware of this, they have a training program looking at personalized medicine and actually have had conversation with some of the folks in that program.

I think that we can -- yes, automate what can be automated but ultimately I do think we're going to have to have that skill set. We're going to have to have that professional skill set to take this information, analyze it and help people stay up to date. I just think it's going to be a huge need and frankly an opportunity.

DR. LITWACK: So we've just got a few minutes left and I want to give the opportunity to the audience to ask questions so if you will come up to the mic and identify yourself, where you're from please.

DR. LINDEMAN: Sure, Neal Lindeman, Brigham Women's Hospital. In the discussion on data sharing there was a lot of discussion about how to incentivize various parties.

There was one party I didn't hear mentioned and that party to me is the one that actually has the most valuable data of all which is the information about the variants that haven't been published ad nauseam.

So how do we incentivize the data that's being investigated and sort of sequestered in research trials and projects and we don't yet know what it means but there might be patients that have similar alterations and I'll sit down.

DR. DICKSON: I mean I think -- so you're talking clinical trial patients?

DR. LINDEMAN: Yeah.

DR. DICKSON: I mean I think there's a lot of push to get that data out of those clinical trials. I mean the FDA has been working on saying what else is there?

There are other private or public groups that have been working on -- Project Data Sphere has been going through and saying look, let's try to collect data and make it available.

Neal I think the question is -- what data aren't being reported and what do we do to get those data upfront so that we can then analyze them later on?

I think as we start to aggregate those data that either comes through the FDA efforts -- it's gone through an NDA or a trial, you know, that's been applied to the FDA or something else -- we are going to have to decide what would we have been able to do if we would have had the following datasets?

And so being the little bit of -- planning upfront a little bit on what we can do with certain data points maybe very valuable but at least if we can start by collecting what we have already collected it's a starting point.

DR. LINDEMAN: And then I'd add the failed trial data may be even more valuable than the successful trial data.

DR. LICHTENFELD: David, can I address that -- good because we obviously have the history -- what happened within the clinical trials and their required to publish, you know, failed data in the government supported trials, right?

So that model already exists. And I will share that some organizations who help support research are asking the question whether or not the data should be after a certain period of time be in the public domain.

So it's not a topic that doesn't have precedent or doesn't have people thinking about it and it's an important one -- positive or negative for that matter.

I think the negative is also something we need to know about.

DR. DEAN: Something that comes to mind about this is an earlier comment is clinical trials are big in lots of different institutions and I wonder if we can break the problem down into personalized collections that could work together where there is added value?

So I think we always want to break it down smaller into sub-groups would be my first answer. But it's still open -- it's a hard one, thank you.

DR. ROSCOE: Hi, thank you for those talks they were really interesting. Um, so the reason -- one of the big reasons why we had this workshop was to try and get at what is meaningful versus what is still in research stages.

And I thought that your talks were really enlightening about what's coming down the road before we even probably manage to address this issue because right now we're all worried, you know, as Dr. Shaw said, you know, should we treat V600L or V600E and what is the anecdotal evidence in patients?

But eventually with more and more data you're going to get people and private enterprises, developing algorithms that incorporate not just that one somatic mutation but germline mutations and sort of other things and that will not be transparent.

We're not going to have the skillset -- we're not going to be able to look at what was the evidence behind that. And it's probably not a leap to see that eventually as you develop more and more proprietary, commercialized enterprises that make use of these databases and develop complicated algorithms it's going to tax the system in terms of attempting to find options -- treatment options for people.

And so it might not be long before we actually see the reverse happening. Well yes, you had this mutation identified but private payers are saying you also have these other mutations so we're saying you're not a responder, you should not be treated because our algorithms, our database have for their own purposes decided that treatments are not going to work for you.

So does anyone have a comment about really helping us and helping the community know how to define the threshold of evidence for what moves forward, what doesn't move forward and also how that does -- that gets used appropriately?

DR. DICKSON: When I talked about quality improvement registries -- QIR's. QIR's are multi-stakeholder organizations that come together to say what do we need to collect? How are we going to collect it? What data are we receiving?

How is that data being used and how do we then improve that data? The problem that we run into when we look just completely retrospectively as we're getting what people thought were important -- when we start looking prospectively and we look at iteratively, there's no reason we can't come together and say okay, we thought this now six months from now what are we going to think differently, particularly with variant calls.

And if we go through and collect a granular genomic dataset -- let's say you collect a BAM file, you collect a BED file and you collect variant calls and maybe you collect them even before curation.

Maybe your concern is with the curation step. If you collect all those data elements and you start seeing that a physician acts in a certain way and a patient doesn't respond to that, you at least have got your individual data points that you can come back and say maybe the problem was not with the testing at all, maybe it was the physician didn't understand the report.

Maybe the problem was that, you know, that the manual curation called something that another group would have said no, it's noise in the background. I mean we don't know where the problems are going to be but it requires that we have infrastructure to build these quality improvement registries that can continually update the testing.

And that's the danger right now of building just, you know, retrospective datasets or building traditional registries where basically they're looking at one test, one treatment, one outcome and reporting it in five years.

We need to be able to look at the data and look at the data -- you know, all of us need to look at the data together and work to decide how is this improving the care of patients and what do we need to change six months from now with version 2.0 which we don't know what it is right now.

DR. LITWACK: And so just in the interest of time because we're running over and somebody has been waiting patiently, one more question.

UNIDENTIFIED SPEAKER: Very nice analogy on the tellers and the ATM. So given the fact that you know, at the end of the day, we're in a situation where the devil is in the details in terms of -- someone has to actually read the papers and understand the evidence and ask some tough questions whether controls were appropriate, whether the study population was appropriate, the data presented is -- does it make sense?

What are your thoughts on natural language processing and artificial intelligent tools to help us with a bibliographic content curation? Now I've seen situations where Google Scholar has picked up a variant in a paper and the authors of the paper had made a type at the nucleotide level for that variant and Google Scholar exactly picked up the type and it made it look like that odd variant was not in the paper, it was a different variant.

So someone actually had to read the paper to realize that the authors had made a typo. This was an outcome of Google Scholar which I think uses certain sophisticated AI tools.

Even to go to a clinical trial and actually have to look at the inclusion criteria and understand the criteria takes time. So again, that would be a big benefit to use AI tools and where do you think the field is going because that would really help save valuable time in evidence curation?

DR. GROSSMAN: Briefly, I talked a little bit about this. I think that's going to be an important technology but the advances we've made in statistical learning that give us Alexa and all that has been not from the deep models on sort of curated data like that and published data and languages but at scale doing simple statistics on all the data.

So I think at scale we're going to be able to do simple things that are going to outperform that over time. So, but that's two cents I mean I've been in the field long enough so whatever you say, the pendulum swings back and we do complex things again.

But right now the pendulum is forcing basic -- is favoring simpler statistics at larger -- that will correct the sort of the bibliographic things, but I'm sure it's going to swing back the other way at one point.

DR. LITWACK: And so we're over so we'll just leave it at that and I want to give a hand to our panel -- that was a great discussion. Thank you very much.

DR. MADISON: Alright thank you panel for that wonderful session and thank you everyone for staying for today and listening to all the talks and those who are online.

We wanted to end the day with a summary because we thought that it would be very important to sort of bring all of these topics and these sessions and the discussions that were taking place today in a context of a big picture look.

And so we have Dr. Schuck from the Center for Devices -- from CDER, not from Center for Devices, from CDER, the Center for Drugs is a clinical pharmacology so he's going to give us our closing remarks and summary of the day.

DR. SCHUCK: Alright everyone so I'm going to try to summarize the last 8 or so hours of in depth scientific discussions in exactly 4 slides so I'll have us out of here in no time.

Before I get started actually I wanted to say thank you to everybody for coming today. I think it's been very enlightening to hear these opinions and perspectives from a very broad group of individuals representing many different institutions across the country as well as many different backgrounds from treating clinicians to data scientists.

So thank you all for coming and helping us out today. I think it's been a very productive conversation. It's really helped us kind of get a good understanding of what the current state of the science is as well as some of the future challenges that are going to be upcoming and how we can help potentially get out in front of those, so thank you.

Also thank you to Hisani and the rest of the organizing committee for putting this together.

Alright so we started out this morning by getting an overview of the state of the science for variant classification and its practice use in treating patients.

So the current state of the science I think could be described -- what I took from that conversation was that there's very many -- a lot of institutions are using these large sequencing panels from everywhere from community hospitals to state of the art academic centers and everywhere in between.

And these are being used to guide patient treatment and enrollment into clinical trials. So interpretation currently relies on using data from multiple different sources and the clinical guidelines don't actually cover all of the scenarios so there are new variants and conflicting data are often being discovered that create new challenges.

So some highlights from the discussion would be that precision oncology is a rapidly evolving field and we need to anticipate changes and think towards the future.

Some of the things that were brought up were moving away from the one drug/one test model towards the use of these panels and also developing frameworks for interpretation that are dynamic in recognizing that the data changes very, very quickly and that we need to not have static programs but have um, frameworks that can move with the field.

In the second panel we discussed levels of evidence that are required for reporting variants and guiding patient treatment. So the current state is that last year the ASP ASCO CAP published a consensus guideline in January to provide a framework for interpretation and reporting of data from sequencing panels.

Some of the highlights that we saw that were discussed were the variant interpretation and that reporting should be standardized as much as possible but the standardization process does not supersede the practice of medicine so there's still some need for a human element in there.

Things that were discussed were the manual interpretation by experts with a broad background can supplement guidelines. In particular cases where this might be useful are in the interpretation of lower tier evidence so perhaps the very common variants with a lot of data behind them are all going to be interpreted the same.

But when it gets down to situations where there's less certainty surrounding that variant it can be helpful to have manual curation there. Also a use of lower levels of evidence as a tie-breaker when multiple therapies are available -- so this got into -- and when I refer to lower levels of evidence here I'm referring to things such as not clinical trial data, but perhaps case series or non-clinical data can help break that tie and help guide patient care in that situation.

Lastly, understanding of test limitations and differences between tests is important and also was perhaps identified as an area where we need to be training clinicians better and where there might be a gap in the field.

That also just came up again in the last session a few minutes ago that this is perhaps the clinical understanding of what the test is actually telling you and what the sensitivity and specificity and specifications of it are is perhaps lacking on that end.

Some next steps that were identified are innovative regulatory strategies that are needed to capitalize on data from sequencing panels and also integration of the panel data with the EMR may facilitate the use of data. I think I'm missing a bullet there.

On topic 3 we moved on to best practices for use of public private databases in variant classification and interpretation in oncology.

The current state was that there are multiple public and private databases to aid in the various classifications and interpretation -- each have their advantages and disadvantages and interpretation across different databases might not always be consistent for a variety of reasons that were discussed.

Some of the discussions that happened during the panel were that transparency and data sources, methods and rules and reporting is important. That was kind of identified as the key factor because we need to be able to understand how the decision to classify was made and how the -- whoever was interpreting it got to that decision point.

So interpretation of functional data and literature sources are often sources of discrepancies um, and this is due to them potentially being outdated or things changing over time and different interpretation of the importance of non-clinical data sources.

And continually updating classifications is challenging but necessary to appropriately care for patients and advance the science.

In our last session we discussed future directions for data sharing, standardization and establishing consistency in precision oncology.

The current state is that although the large scale sharing of data is very difficult, there are certainly some great efforts underway to create large databases of genomic and clinical data that can create useful information for the treatment of patients.

Some highlights of the discussion are that a more flexible regulatory payer and healthcare systems are necessary to advance precision medicine.

We need to develop metrics to ensure the quality of the data and also building infrastructure and training programs that facilitate the appropriate use and analysis of large datasets will be key to moving this field forward.

And the next steps are to create an environment that facilitates the generation of useful, large scale databases for patients. So with that again I’d like to on behalf of the organizing committee say thank you all very much for joining us today and Hisani did you want to add anything to wrap it up -- alright, thank you all very much.

CERTIFICATE OF NOTARY PUBLIC

I, NATALIA THOMAS, the officer before whom the foregoing proceeding was taken, do hereby certify that the proceedings were recorded by me and thereafter reduced to typewriting under my direction; that said proceedings are a true and accurate record to the best of my knowledge, skills, and ability; that I am neither counsel for, related to, nor employed by any of the parties to the action in which this was taken; and, further, that I am not a relative or employee of any counsel or attorney employed by the parties hereto, nor financially or otherwise interested in the outcome of this action.

Natalia Thomas

NATALIA THOMAS

Notary Public in and for the

State of Maryland

CERTIFICATE OF TRANSCRIBER

I, HELEN VENTURINI, do hereby certify that this transcript was prepared from audio to the best of my ability.

I am neither counsel for, related to, nor employed by any of the parties to this action, nor financially or otherwise interested in the outcome of this action.

February 6, 2018 Helen Venturini

DATE HELEN VENTURINI